

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2005, 02:19:48 ; Search time 165 Seconds  
(without alignments)  
673.188 Million cell updates/sec

Title: US-10-657-852B-15

Perfect score: 1385

Sequence: 1 MPEYMAKCCMLLVFLGFIQ.....NTVSGSNHIVSGSNKVVTDG 267

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1854112 seqs, 416015017 residues

Total number of hits satisfying chosen parameters: 1854112

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1385	100.0	267	18	US-10-657-852-16
2	1290	93.1	269	18	US-10-657-852-17
3	992.5	71.7	262	18	US-10-657-852-24
4	980.5	70.8	254	18	US-10-657-852-23
5	973.5	70.3	256	18	US-10-657-852-25
6	910.5	65.7	277	18	US-10-657-852-20
7	908.5	65.6	277	18	US-10-657-852-19
8	904.5	65.3	281	18	US-10-657-852-26
9	901.5	65.1	281	18	US-10-657-852-18
10	901.5	65.1	285	18	US-10-657-852-22
11	884	63.8	280	18	US-10-657-852-21

12	668	48.2	243	18	US-10-657-852-15	Sequence 15, Appl
13	655	47.3	243	18	US-10-657-852-14	Sequence 14, Appl
14	262	18.9	1010	16	US-10-437-963-180288	Sequence 180288
15	206.5	14.9	1039	16	US-10-437-963-16082	Sequence 16082
16	204	14.7	1021	15	US-10-443-101-2	Sequence 2, Appli
17	204	14.7	1021	17	US-10-733-923-1223	Sequence 1223, Ap
18	198	14.3	264	16	US-10-437-963-148813	Sequence 148813
19	197	14.2	961	16	US-10-437-963-172660	Sequence 172660
20	194	14.0	917	16	US-10-437-963-115627	Sequence 115627
21	194	14.0	1164	16	US-10-437-963-192784	Sequence 192784
22	192.5	13.9	838	16	US-10-437-963-169754	Sequence 169754
23	192	13.9	1060	16	US-10-437-963-153194	Sequence 153194
24	190	13.7	1044	16	US-10-437-963-162505	Sequence 162505
25	188.5	13.6	1011	16	US-10-437-963-172329	Sequence 172329
26	188.5	13.6	1078	16	US-10-437-963-123119	Sequence 123119
27	186	13.4	278	15	US-10-424-599-267766	Sequence 267766
28	186	13.4	676	16	US-10-437-963-153306	Sequence 153306
29	185	13.4	1012	16	US-10-437-963-192786	Sequence 192786
30	183.5	13.2	380	16	US-10-437-963-136619	Sequence 136619
31	183.5	13.2	1133	16	US-10-437-963-150876	Sequence 150876
32	183.5	13.2	1140	16	US-10-437-963-119802	Sequence 119802
33	182	13.1	941	10	US-09-952-267-9	Sequence 9, Appli
34	182	13.1	941	18	US-10-872-768-9	Sequence 9, Appli
35	182	13.1	941	18	US-10-872-769-9	Sequence 9, Appli
36	182	13.1	1231	16	US-10-437-963-102913	Sequence 102913
37	181	13.1	653	15	US-10-389-566-1625	Sequence 1625, Ap
38	181	13.1	1065	16	US-10-437-963-109994	Sequence 109994
39	181	13.1	1469	16	US-10-437-963-196963	Sequence 196963
40	180.5	13.0	675	16	US-10-437-963-174437	Sequence 174437
41	180	13.0	1036	16	US-10-437-963-149679	Sequence 149679
42	180	13.0	1080	16	US-10-437-963-110291	Sequence 110291
43	179	12.9	516	15	US-10-424-599-201138	Sequence 201138
44	178	12.9	143	16	US-10-425-115-295330	Sequence 295330
45	178	12.9	153	15	US-10-425-114-63986	Sequence 63986, A

ALIGNMENTS

RESULT 1  
US-10-657-852-16  
; Sequence 16, Application US/10657852  
; Publication No. US20040146884A1  
; GENERAL INFORMATION:  
; APPLICANT: Demmer, Jercoen  
; APPLICANT: Shenk, Michael Andrew  
; APPLICANT: Hall, Claire  
; APPLICANT: Fish, Steven A  
; TITLE OF INVENTION: Antifreeze proteins isolated from forage  
; FILE OF INVENTION: Grasses and methods for their use.  
; FILE REFERENCE: 11000.1070U  
; CURRENT APPLICATION NUMBER: US/10/657,852  
; CURRENT FILING DATE: 2003-09-09  
; PRIOR APPLICATION NUMBER: 60/409,557  
; PRIOR FILING DATE: 2002-09-09  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Festuca arundinacea  
US-10-657-852-16

Query Match	100.0%	Score 1385;	DB 18;	Length 267;
Best Local Similarity	100.0%;	Pred. No. 1.9e-111;		
Matches 267;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPEYMAKCCMLLVFLGFIQVAGATSWCHDDHLALGLAENLSGKGVRLRAAWSGAS	60	
Db	1	MPEYMAKCCMLLVFLGFIQVAGATSWCHDDHLALGLAENLSGKGVRLRAAWSGAS	60	
QY	61	CCSWEVGVCETASGRVVALRLPKRGLGGIIPSGISGLDHLRYLDLSNSLVGVPSKLOI	120	

```

Db 61 CCSWEGVGCETASGRVVALRLPKRGLGGIIPSSIGELDHLRLYLDLGSNSLVGEVPKSLQI 120
QY 121 RLKSLTTDSQSLGMSINMLLHVSSRRRLTDEEPNTISGTNNVSGSGNNVSGNDNTVVS 180
Db 121 RLKSLTTDSQSLGMSINMLLHVSSRRRLTDEEPNTISGTNNVSGSGNNVSGNDNTVVS 180
QY 181 GNNNHVSGSNNTVVTGSDNTVVGSHVSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEH 240
Db 181 GNNNHVSGSNNTVVTGSDNTVVGSHVSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEH 240
QY 241 NTVSGSNNTVSGSNHIVSGSNKVVTDG 267
Db 241 NTVSGSNNTVSGSNHIVSGSNKVVTDG 267

```

RESULT 2

```

US-10-657-852-17
; Sequence 17, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Festuca arundinacea
; NAME/KEY: VARIANT
; LOCATION: (1)...(269)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-657-852-17

```

```

Query Match 93.1%; Score 1290; DB 18; Length 269;
Best Local Similarity 93.3%; Pred. No. 3.1e-103;
Matches 251; Conservative 6; Mismatches 10; Indels 2; Gaps 2;

QY 1 MPEYMAKCCMLLVFLGFI-LQVAGATWSCHHDDLHALRGLAENLSGKGAVALRLAAWSGA 59
Db 1 MPEYMAKCCMLLVALLAFILLQVAGATWSCHHDDLALRGLAENLSGKGAVALRLAAWSGA 60
QY 60 CCSWEGVGCETASGRVVALRLPKRGLGGIIPSSIGELDHLRLYLDLGSNSLVGEVPKSLQ 119
Db 61 CCSWEGVGCETASGRVVALRLPKRGLGGIIPSSIGELDHLRLYLDLGSNSLVGEVPKSLQ 120
QY 120 IRLKSLTTDSQSLGMSINMLLHVSS-RRTLDEEPNTISGTNNVSGSGNNVSGNDNTV 178
Db 121 IRLKSLTTDSQSLGMSINMLLHVSSRRRLTDEEPNTISGTNNVSGSGNNVSGNDNTV 180
QY 179 VSGNNNHVSGSNNTVVTGSDNTVVGSHVSGTKHIVTDNNNVSGNDNNVSGSFHTVSG 238
Db 181 ISGNNNHVSGSNNTVVTGSDNTLVGSHVSGTKHIVTDNNNVSGNDNNVSGSFHTVSG 240
QY 239 EHNVTSGSNNTVSGSNHIVSGSNKVVTDG 267
Db 241 EHNVTSGSNNTVSGSNHIVSGSNKVVTDG 269

```

RESULT 3

```

US-10-657-852-24
; Sequence 24, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-657-852-23

```

```

; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-657-852-24

```

```

Query Match 71.7%; Score 992.5; DB 18; Length 262;
Best Local Similarity 75.0%; Pred. No. 1.6e-77;
Matches 198; Conservative 22; Mismatches 39; Indels 5; Gaps 3;

QY 5 MAKCCMLLVFLGFILOVAG-ATWSCHHDDLHALRGLAENLSGKGAVALRLAAWSGASCCS 63
Db 1 MAKCLMLLLSFAFLLSAAGTATATPCHRDDLRALRGLAENLSGKGAVALRLAAWSGASCCD 60
QY 64 WEGVGCETASGRVVALRLPKRGLGGIIPSSIGELDHLRLYLDLGSNSLVGEVPKSLQIRLK 123
Db 61 WEGVGCDSGSGRVVALRPSGLTGPIPSWICQLHRLYLDLGSNALVGEVPKNLQVQLK 120
QY 124 SLTTDSQSLGMSINMLLHV-SSRTLDEEPNTISGTNNVSGSGNNVSGNDNTVVS 182
Db 121 GLTAAGRS---GFTNNPLHVNRRLSRLDEQPNITSGSNNTVRSGKNVVGAGNDNTVISGD 177
QY 183 NNNHVSGSNNTVVTGSDNTVVGSHVSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEHN 242
Db 178 NNSVSGSNNTVVTGSDNTVVGSHVSGTNHIVTDNNNVSGNDNNVSGSFHTVSGCHNT 237
QY 243 VSGSNNTVSGSNHIVSGSNKVVTD 266
Db 238 VSGSNNTVSGSNHIVSGSNKVVTD 261

```

RESULT 4

```

US-10-657-852-23
; Sequence 23, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-657-852-23

```

```

Query Match 70.8%; Score 980.5; DB 18; Length 254;
Best Local Similarity 73.5%; Pred. No. 1.7e-76;
Matches 194; Conservative 24; Mismatches 33; Indels 13; Gaps 3;

QY 5 MAKCCMLLVFLGFILOVAG-ATWSCHHDDLHALRGLAENLSGKGAVALRLAAWSGASCCS 63

```

[illegible]

```

RESULT 5
US-10-657-852-25
; Sequence 25, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; FILE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-657-852-25

```

RESULT 6  
US-10-657-852-20  
; Sequence 20, Application US/10657852  
; Publication NO. US20040146884A1

```

; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Festuca arundinacea
; US-10-657-852-20

```

```

RESULT 7
US-10-657-852-19
; Sequence 19, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use.
; FILE REFERENCE: 11000.1070U
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-657-852-19

```









```

QY      214   IVTDNNNVSGNDNNVSGSFHTVSGETWTS-----GSNNTVSGSNHIVSG---SN 261
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      164   QANTDNAVGLKNTTNGNNSAAIGSENTVNENKNVFILGNTTNQAQSGSVLLGHETSG 223

QY      262   KVVVT 265
       || | |
Db      224   KEAT 227

RESULT 15
US-10-101-464A-732
; Sequence 732, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 732
; LENGTH: 187
; TYPE: PRF
; ORGANISM: Pinus radiata
US-10-101-464A-732
```

```
Db 119 LQELNLASNGLS-GTIPYSISNMTGLTDLKLSHNQLSGQIQIDIFGQLSSLSLTLDLSFNTL 177
Qy 157 SGT-----NNSVSGSNVNVSGNDNTVVSGNNHVS----- 188
Db 178 TQNLPOSFSSLSLVLYLQNNQL-AGSVNVLANPLTDLNENRFSGWNPMNRSNQN 236
Qy 189 ----SNNVTVG-----SDNTVVGSHVV---SGPKHIVTDNNNVVSGNDNN 228
Db 237 FKYSNGSFATGAPPPPPPTPPPPNNRPPKSSNVVPSGGSK-----GGSNK 286
Qy 229 VSGSFHTVSG 238
Db 287 KSLSGGAIVG 296

RESULT 12
US-10-101-464A-80
; Sequence 80, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-80

Query Match 12.3%; Score 170.5; DB 4; Length 707;
Best Local Similarity 26.1%; Pred. No. 8.9e-08;
Matches 81; Conservative 39; Mismatches 89; Indels 101; Gaps 16;

Qy 9 CMLLVFLGFILOVA---GATSNCHDDLHLRGLAENLSGKGAVRLRAWS--GASCC- 62
Db 8 CURLIAILTASTAISGHGTTD----PDVSAKGIYSLNSPOOL---SGWSANGDPCG 60

Qy 63 -SWEGVGCETASGRVVALRPKRGLGIIIPSSIGELDLRLYLDLSGNSLVGEVPKSLQIR 121
Db 61 QSWKGVCSCSS--VTLIKSLGSLGSLYYQLSDLSSTLTLDLSNNIQNPYALPKQ 118

Qy 122 LKSLTTDSQSLMGSI-----NM-----LH-----VSSRRITLDEEPTI 156
Db 119 LQELNLASNGLS-GTIPYSISNMTGLTDLKLSHNQLSGQIQIDIFGQLSSLSLTLDLSFNTL 177

Qy 157 SGT-----NNSVSGSNVNVSGNDNTVVSGNNHVS----- 188
Db 178 TQNLPOSFSSLSLVLYLQNNQL-AGSVNVLANPLTDLNENRFSGWNPMNRSNQN 236
Qy 189 ----SNNVTVG-----SDNTVVGSHVV---SGPKHIVTDNNNVVSGNDNN 228
Db 237 FKYSNGSFATGAPPPPPPTPPPPNNRPPKSSNVVPSGGSK-----GGSNK 286
Qy 229 VSGSFHTVSG 238
Db 287 KSLSGGAIVG 296
```

```
RESULT 13
US-09-336-447A-7
; Sequence 7, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-7

Query Match 12.3%; Score 170; DB 3; Length 624;
Best Local Similarity 30.6%; Pred. No. 8.3e-08;
Matches 38; Conservative 24; Mismatches 50; Indels 12; Gaps 2;

Qy 154 NTISGTNNSVSGSNVNVSGNDNTVVSGNNHVSNNNTVVTGSDNTVVSGNNHVSGETKH 213
Db 104 NQAKGEHSTIAGSGNQATGRNSTVAGGSNNQAVGTNSTVAGGSNNQAKGANSFAAGVGN 163

Qy 214 IVTDNNNVSGNDNVSGSFHTVSGEHTVS-----GSDNTVSGSNHIVSG---SN 261
Db 164 QANTDNAVGKNTINGNNSAAIGSENTVNNQKNVFILGNTTNAQSGVLLGHETSG 223

Qy 262 KVVVT 265
Db 224 KEAT 227

RESULT 14
US-09-952-267B-7
; Sequence 7, Application US/09952267B
; Patent No. 6753417
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/952,267B
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/336,447
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-7

Query Match 12.3%; Score 170; DB 4; Length 624;
Best Local Similarity 30.6%; Pred. No. 8.3e-08;
Matches 38; Conservative 24; Mismatches 50; Indels 12; Gaps 2;

Qy 154 NTISGTNNSVSGSNVNVSGNDNTVVSGNNHVSNNNTVVTGSDNTVVSGNNHVSGETKH 213
Db 104 NQAKGEHSTIAGSGNQATGRNSTVAGGSNNQAVGTNSTVAGGSNNQAKGANSFAAGVGN 163
```





```
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/336,447
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 941
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-9
```

```
Query Match      13.1%; Score 182; DB 4; Length 941;
Best Local Similarity 34.2%; Pred. No. 1e-08;
Matches 39; Conservative 19; Mismatches 50; Indels 6; Gaps 1;

QY 154 NTISGTNNVSGSNVYVSGNDNTVVGNNHVSNNNTVVTGSDNTVVGSNHVSNGTKH 213
Db 123 NEAMGEYSTVAGGANNQAGNYSYTVGGNGNKAIGNNSTVVGSNQAGHSTIAGGK- 181

QY 214 IVTDNNVSGNDNNVSGFHTVSGEHTVSGNNNTVSGSNHVSNNKVVTDG 267
Db 182 -----NQATGGSFAAGVENKADANNAVALGNKNTIEGTNSVAIGSNNTVKTG 230
```

```
RESULT 3
US-09-336-447A-5
; Sequence 5, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: ABEI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-5
```

```
Query Match      12.8%; Score 177; DB 3; Length 892;
Best Local Similarity 31.8%; Pred. No. 2.9e-08;
Matches 35; Conservative 20; Mismatches 49; Indels 6; Gaps 1;

QY 158 GTNNVSGSNVYVSGNDNTVVGNNHVSNNNTVVTGSDNTVVGSNHVSNGTKHIVTD 217
Db 143 GDSSTIGGGYNNQATGEKSTVAGGRNNQATGNNSTVAGGSYNQATGNNSTVAGGSH---- 198

QY 218 NNNVSGNDNNVSGFHTVSGEHTVSGNNNTVVGSNHVSNNKVVTDG 267
Db 199 --NQATGGSFAAGVENKANANNAVALGNKNTIDGNSVAIGSNNTIDSG 246
```

```
RESULT 4
US-09-952-267B-5
; Sequence 5, Application US/09952267B
; Patent No. 6753417
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: ABEI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
```

```
; CURRENT APPLICATION NUMBER: US/09/952,267B
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/336,447
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-5
```

```
Query Match      12.8%; Score 177; DB 4; Length 892;
Best Local Similarity 31.8%; Pred. No. 2.9e-08;
Matches 35; Conservative 20; Mismatches 49; Indels 6; Gaps 1;

QY 158 GTNNVSGSNVYVSGNDNTVVGNNHVSNNNTVVTGSDNTVVGSNHVSNGTKHIVTD 217
Db 143 GDSSTIGGGYNNQATGEKSTVAGGRNNQATGNNSTVAGGSYNQATGNNSTVAGGSH---- 198

QY 218 NNNVSGNDNNVSGFHTVSGEHTVSGNNNTVVGSNHVSNNKVVTDG 267
Db 199 --NQATGGSFAAGVENKANANNAVALGNKNTIDGNSVAIGSNNTIDSG 246
```

```
RESULT 5
US-10-101-464A-890
; Sequence 890, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 890
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-890
```

```
Query Match      12.7%; Score 176.5; DB 4; Length 804;
Best Local Similarity 23.4%; Pred. No. 2.8e-08;
Matches 77; Conservative 54; Mismatches 105; Indels 93; Gaps 13;

QY 10 MLLVFLGFL-----QVAGATSWCHD--DLHALRGLAE--NL 44
Db 6 LLLVMGVAMPTHQSQTGFTSVQRPFPNGRSMGKPSIAGYHEKRDVEALLSFRKGITL 65

QY 45 SGKAVRLRAAASGASCCSWEGVGETAGRVVALRLPKRGLGGIIPSGIGDLHLRYLD 104
Db 66 DPYGLSNWTANNHSHNVCIMNGISCPNTNRVVEISLRYGRNLGTLSPYIGNLSLRLHD 125

QY 105 LSGNSILVEVPKSL----QIRLKSITDTSQS-----LGMGSINMLLHVSRRRTLDEEPN 154
Db 126 LSSNALSGRIPAEFGQLKALRIIDLSNNALTGSIPTCIGNGGINGTL-----DLDSWN 179

QY 155 TTSG-----TNNSVSGSNVSGNDNTVVGNN--NHVSGSNNTVV- 194
Db 180 AFSGRIPKELFNCTRLQRLDLSHNSL-TGSIPTSGNCALLQTLNTGPNLVSIGSIPTSLA 238
```





```
Db 229 LTVA--INNLOGLIPVLFNMSLECLNFGSNQSGS 263
RESULT 14
Q66CJ0 PRELIMINARY; PRT; 622 AA.
AC Q66CJ0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Putative exported protein precursor.
GN ORFNames=YPTB1413;
OS Yersinia pseudotuberculosis IP 32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
NCBI_TaxID=273123;
[1]
SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX PubMed=15358858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
Brubaker R.R., Fowler J., Hinnbusch B.J., Marceau M., Medigue C.,
Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
Derbise A., Hauser L.J., Garcia E.;
RA "Insights into the genome evolution of Yersinia pestis through whole
genome comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
KW EMBL; BX936398; CAH20653.1; -.
FT SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1 32 Potential.
Query Match 14.0%; Score 194; DB 2; Length 622;
Best Local Similarity 26.8%; Pred. No. 3.6e-06;
Matches 68; Conservative 48; Mismatches 78; Indels 60; Gaps 12;
QY 43 NLSGKAVRLRAWVGSCCSWEGVCETASGRVVALRPLKRLGGIIPSSIGELDLRLY 102
Db 41 NNNNGTINIFDASSNNDIHTLTGLNEL-----LGGF-----SNHL-- 77
QY 103 LDLSGNSLVGEVPSKQIRLKSITLTDQSLSGMS-----SINMLLVSSRRT-L 149
Db 78 IDSHNTINGQSQNNL-VSSDGNNTISALSLGDLGYGAQNNTLINSNNLLIVTQSTII 136
QY 150 DEEPNTISGTNNVSGSNNVSGNDNTVVS-----GNNHVSNNVTVVTS 198
Db 137 DSDSNTVSGISNNLIESNNII-GNENSCVSDPASPAGWCDNQNTLIGSDNNTITGAL 195
QY 199 NTVVGSNH---VSGTKHIVTDNNNVSGNDNNVSGSFHTVS-GEHNTVSGSNNVTS 254
Db 196 NGLHNSHNDIIASSVNNMLDTHNNIAG-----GHYNTISGGNNDIFGSNNVTDST 249
QY 255 HI-VSGSNKVVTDG 267
Db 250 DANINGSNNYVIDG 263
[1]
RESULT 15
Q75GM9 PRELIMINARY; PRT; 917 AA.
AC Q75GM9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein OSJNBa0018K15.10.
GN Name=OSJNBa0018K15.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
NCBI_TaxID=39947;
[1]
```

```
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Hsiung Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Lee P.-F., Chang S.-J., Chen H.-C., Chen S.-K.,
RA Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y.,
RA Hsiao S.-H., Hsiung J.-N., Hsu C.-H., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC144737; AAT01367.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR InterPro; IPR003591; LRR_typ.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00560; LRR_1; 14.
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00369; LRR_TYP; 5.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Hypothetical protein.
SQ SEQUENCE 917 AA; 96507 MW; AE6DC9F7DD6E0B77 CRC64;
Query Match 14.0%; Score 194; DB 2; Length 917;
Best Local Similarity 26.3%; Pred. No. 5.9e-06;
Matches 81; Conservative 43; Mismatches 108; Indels 76; Gaps 12;
QY 6 AKCCMLLVPLGFILQVAGATSWCHHDDLHRLRGLAENLSGKGVRLRAWS--GASCCS 63
Db 4 ATAALLLLTAAIAAGAV-----NDPVLALVVFKSGVSDPGGV--LAAWSEADARACA 56
QY 64 WEGVCETASGRVVALRPLKRLGGIIP-SSIGELDLRLYLDLSGNSLVGEVPSKQIRL 122
Db 57 WFGVSCDARAGRVDVAVALPSAGLSGRLPRLSALLRLDALLSLALPGNLSGPIPDALPPRA 116
QY 123 KSLTTTDSOSL-----GMGSINN-----LLHVSRRTLDEEPNTISG 158
Db 117 RALDLSANSLSGYLPAAALASCGSLVSLNLSGNLLSGVPDGIWSLPSLRLSLDSGNQLAG 176
QY 159 TNNSVSGS-----SNNVSGNDNTVVS-----GNNHVSNNVTV--V 194
Db 177 ---SVPGGFPRSSRLRLDLRLNLEGEIPADVGEAGLLKSLDVGHNLFTELPSLRGL 233
QY 195 TGSNTVVSNNH-----VSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEHNTVTS 244
Db 234 TGLSLGAGGNALAGELPGWIGEMAALETLDLSGNRFVGAI PDGISGCKNLVE-----VD 288
QY 245 GSNNTVSG 252
Db 289 LSGNALTG 296
```

Search completed: October 8, 2005, 02:19:41  
Job time : 180 secs









DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Putative outer membrane protein.  
GN ORFNames=BPSL1631;  
OS Burkholderia pseudomallei K96243.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Burkholderia.  
OX NCBI\_TaxID=272560;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K96243;  
RX PubMed=15377794;  
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdono-Tarraga A.M.,  
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,  
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,  
RA Brooks K., Brown K.A., Brown N.P., Challis G.L., Cherevach I.,  
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,  
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,  
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,  
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,  
RA Songvilalai S., Stevens K., Tumapa S., Vesaratchaveat M.,  
RA Whitehead S., Yeats C., Barrall B.G., Oyston P.C.F., Parkhill J.,  
RT "Genomic plasticity of the causative agent of melioidosis,  
RT Burkholderia pseudomallei";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).  
DR EMBL; BX571965; CAH35630.1; -;  
SQ SEQUENCE 1124 AA; 104844 MW; 90EC4527E9017602 CRC64;  
  
Query Match 16.0%; Score 221.5; DB 2; Length 1124;  
Best Local Similarity 27.7%; Pred. No. 6.6e-08;  
Matches 74; Conservative 41; Mismatches 99; Indels 53; Gaps 9;  
  
QY 36 ALRGLAENLSGKAVRLRAAWSGCCSWEGVCETASGRVVALRLPKR-----GLGGI 89  
DB 316 SLIGTADNVTSLS-----TSLSTVNANLAGLQTSVDNVSYDDPSKSAITLGGAGVT 367  
  
QY 90 IPSISIGELHLRLVDLSGNSLVEVPKSLQ-----TLKSLTTDSQSLGMSGIN 138  
DB 368 TPVLLTNVAAGKIAATSTDAVNGSOLYTLQQEFSSQYQYLLTSQVSLSTSVSGL-QGSVS 426  
  
QY 139 MLLHVSRRRLDEEPTISGTN-----NSVSGSNVNVGNDNTVVSGNNHVSNNNT 192  
DB 427 -----ANTGTASGDNSTASGDNATASGTNTANGTNTASG-DNSTASGTNASASGNS 480  
  
QY 193 VVTGSDNTVVGVS-----NHVVSFGTKHVTVDNNNVSGNDNNVSGSFHTVSG 238  
DB 481 -ATGTDSTASGNSANGTNTASGDNSTASGTNASATGNSATGTGTDSTASGNSANG 539  
  
QY 239 EHVTVSGSNVTVSGSNHIVSGSNKVV 265  
DB 540 TNSTASGDNSTASGTNASASGNSAT 566  
  
RESULT 5  
Q62KN5 PRELIMINARY; PRT; 1012 AA.  
AC Q62KN5;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DE Outer membrane protein, putative.  
GN ORFNames=BMA1027;  
OS Burkholderia mallei ATCC 23344.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Burkholderia.  
OX NCBI\_TaxID=243160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 23344;  
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,  
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,  
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,

RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,  
RA Mohammad Y., Nelson W.C., Radune D., Romero C.M., Sarrisa S.,  
RA Selenkoff J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,  
RA Zhou L., Fraser C.M.;  
RT "Structural flexibility in the Burkholderia mallei genome.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).  
DR EMBL; CP000010; AAU48666.1; -;  
SQ SEQUENCE 1012 AA; 94758 MW; 53C27D16C62FBDE CRC64;  
  
Query Match 15.7%; Score 218; DB 2; Length 1012;  
Best Local Similarity 30.9%; Pred. No. 1.1e-07;  
Matches 55; Conservative 31; Mismatches 72; Indels 20; Gaps 4;  
  
QY 106 SGNLSVGEVPKSLQIRLKSITTDTSQSLGMSGINMLLHVSRRRLDEEPTISGTN----- 160  
DB 545 SGGDSSTASGTNASATGNSATGTGTDSTASGNSANGTNS--TASGDNSTASGTNASATG 602  
  
QY 161 -NSVSGSNVNVSGND-----NTVSGNNHVSNNHVSNNHVSNNHVSNNHVSNNHVSNNH 207  
DB 603 ENSTATGTDSTASGNSANGTNTASGDNSTASGTNASATGNSATGTGTDSTASGNSNST 662  
  
QY 208 VSGTKHIVTDNNNVSGNDNNVSGSPHTVSGBEHTVSGNNVTVSGSNHIVSGSNKVV 265  
DB 663 ANGANSTASGNSATGTGTDSTASGNSANGTNTASGDNSTASGTNASATGNSATGNSAT 720  
  
RESULT 6  
Q9ZNX4 PRELIMINARY; PRT; 353 AA.  
AC Q9ZNX4;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE PGPS/D4 precursor (LRR protein S/D4).  
GN Names=PGPS/D4;  
OS Petunia hybrida (Petunia).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamids; Solanales; Solanaceae; Petunia.  
OX NCBI\_TaxID=4102;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Germinating petunia pollen treated with kaempferol;  
RX MEDLINES=20317212; PubMed=10859200; DOI=10.1104/pp.123.2.699;  
RA Guyon V.N., Astwood J.D., Garner E.C., Dunker A.K., Taylor L.P.;  
RT "Isolation and characterization of cDNAs expressed in the early stages  
RT of flavonol-induced pollen germination in petunia.";  
RL Plant Physiol. 123:699-710(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Germinating petunia pollen treated with kaempferol;  
RA Guyon V., Astwood J.D., Taylor L.P.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Guyon V.N., Astwood J.D., Garner E.C., Dunker A.K., Taylor L.P.;  
RT "Isolation and characterization of cDNAs expressed in the early stages  
RT of flavonol-induced pollen germination in petunia.";  
RL Life Sci. Adv. Plant Physiol. 123:699-710(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX PubMed=15272880; DOI=10.1111/j.1365-3113X.2004.02162.x;  
RA Guyon V., Tang W.H., Monti M.M., Raiola A., Lorenzo G.D.,  
RA McCormick S., Taylor L.P.;  
RT "Antisense phenotypes reveal a role for SHY, a pollen-specific  
RT leucine-rich repeat protein, in pollen tube growth.";  
RL Plant J. 39:643-654(2004).  
DR EMBL; AF049920; AAD02546.2; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR007090; LRR\_plant.  
DR Pfam; PF00560; LRR\_1; 8  
DR PRINTS; PR00019; LEURICHRPT.



```

RN RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare (GA3) genomic DNA, chromosome 7, PAC
RL clone:PO585H11.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AP004342; BAC20742.1; -.
DR HSSP; P36937; LIAS.
DR Gramene; Q8H3W8; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007030; LRR plant.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00560; LRR_1; 20.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 1010 AA; 109074 MW; 5CCB04E51E18AF02 CRC64;

Query Match 18.9%; Score 262; DB 2; Length 1010;
Best Local Similarity 40.1%; Pred. No. 5.5e-11;
Matches 71; Conservative 19; Mismatches 53; Indels 34; Gaps 5;

QY 7 KCML-LVFLGFTLVAGATSWCHDDHLALRGLAENLS-GKGAVALRAAASGCCSW 64
DB 8 RCLFSLVALFALLPPPPAAAAPCHPEDLLAURAFAGNUSAGGGAGLRAASGDRACCAW 67
QY 65 EGVGCEATAGRVVALRPRGLGIIPSSIGELDHLRYLDLSGNSIVGVPKSLQIRLKS 124
DB 68 DGVACD-AAARVVALRPRGLGIIPSSLAALRLQDLDSHNAIT----- 113
QY 125 LTTDSQSLGWSINMLLVSSRRTLDEEPTISGT-----NNSVSGSNNVSG 173
DB 114 -----GGISALLAAVSLRTANLSSNLLNDTLDLAALPHLSAFNASNLSGG 160

RESULT 3
Q9FN37 PRELIMINARY; PRT; 1036 AA.
AC Q9FN37
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative receptor protein kinase.
GN Name=At5g53890;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RL physically assigned P1 clones.";
RN DNA Res. 4:401-414 (1997).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB007644; BAB10719.1; -.
DR EMBL; AY064019; AAL36375.1; -.
DR EMBL; AY091180; AAL14119.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007030; LRR plant.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00560; LRR_1; 19.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 1036 AA; 114339 MW; 6DF9511FC2A4E261 CRC64;

Query Match 18.7%; Score 258.5; DB 2; Length 1036;
Best Local Similarity 30.7%; Pred. No. 1e-10;
Matches 89; Conservative 50; Mismatches 98; Indels 53; Gaps 14;

QY 10 MLLVFLGFTLVAGATSWCHDDHLALRGLAENLSGKGAVALRAAAW-SGASCCSWEGVG 68
DB 5 LLLVFF-----VGSSVSQFCHPNDLSALRELAKNKSQVTE---SWLNGSRCCWDGVF 56
QY 69 CE--TASGRVALRPRGLGIIPSSIGELDHLRYLDLSGNSIVGVPKSL-----QIRL 122
DB 57 CEGSDVSGRVTKLVLPKGLGVISKIGELTRELRLVLDLSRNQKGEVPAEIKLEQLQV 116
QY 123 KSLTDSQSLGWSI-NMLLVSSRRTLDEEPTISGTNNVSG-----SGSNNVSG 173
DB 117 LDL---SHNLLSGVLGVVSGKLQSLNISNSGKLSGKLSDVGFPLVNLVSNLFEF 173
QY 174 NNTVV---SGNNNHVSGSNTVVTGSDNTVVGSHNVHSGTKHIVTDNNNVWSG----- 224
DB 174 EIHPELCSSGGIGQVLDLNNRLVGNLD-----GLYNCYSKIQQLHDSNRLTQQLPDYLY 229
QY 225 -----NDNNVSGSFHTVSGEHNVTSGSNNVTSGSNNHVSQSNK---VVTD 266
DB 230 SIRELQSLSGNY--LSGE---LSKNLSNLSGLKSLISNRFSDVIPD 274

RESULT 4
Q63UH1 PRELIMINARY; PRT; 1124 AA.
ID Q63UH1
AC Q63UH1;

```



A:Cross-references: UNIPROT:Q22178; EMBL:AC002391; NID:G2642427; PID:G2642433  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: A84623  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-773 <STO>  
A:Cross-references: GB:AE002093; NID:G2642433; PIDN:AAB87101.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g23300; T20D16.7  
A:Map position: 2  
A:Introns: 545/1  
C:Superfamily: protein kinase homology

Query Match 11.7%; Score 162; DB 2; Length 773;  
Best Local Similarity 27.3%; Pred. No. 0.00028;  
Matches 66; Conservative 36; Mismatches 86; Indels 54; Gaps 8;

QY 26 SMSCHDDHLALRGLAENLSGKAVRLRAAWGASCCSWEGVCETASGRVVALRLPKRG 85  
DB 55 SNNYDHD-----NPCSRRGVLNNDSS-RVVTLSLPNSN 86

QY 86 LGGIIPSSIGELDHLRYLDLSGNLSLVEVPKSL-----QIRLSLTDSQS-----LGMGSI 137  
DB 87 LVGSPDLGFLQNLQSLNLSNLSLGPVEFFAADKLRLDLSNLSIGELPVSIGL 146

QY 138 NMLLVHSSRRTLDEEPNTISGTNNVSGSNVNDNTVSGNNHVVSGNNVTVTGS 197  
DB 147 HNL-----QTLNLSNITFTKLPA-----NLASLSLSEVSLKNYFSGEPFGGGRS 194

QY 198 DNTVGSNNHVVST--KHIVTDNNVSGNDNNVSGSFHTVG-----EHTVSGSNNTVS 251  
DB 195 VQYLDISSNLNLSLPPDFSGDNLRYLNVSYNQISGEIPNPVGVAGFPQATVDFSNLTL 254

QY 252 GS 253  
DB 255 GS 256

RESULT 15  
T00475  
N:Alternate names: hypothetical protein F19I3.16  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T00475; F84762  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, April 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.  
A:Reference number: Z14160  
A:Accession: T00475  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-905 <ROU>  
A:Cross-references: UNIPROT:O64757; EMBL:AC004238; NID:G3033373; PID:G3033389  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84762  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-905 <STO>  
A:Cross-references: GB:AE002093; NID:G3033389; PIDN:AAC12833.1; GSPDB:GN00139  
C:Genetics:

A:Gene: F19I3.16; At2g34930  
A:Map position: 2

Query Match 11.7%; Score 161.5; DB 2; Length 905;  
Best Local Similarity 23.1%; Pred. No. 0.00037;  
Matches 88; Conservative 45; Mismatches 105; Indels 143; Gaps 16;

QY 10 MLLVFLGFLQVAGATNSCHDDHLAL-----RGLAENLSGKAVRLRAAWGASCCSWEG 66  
DB 14 LILILLKNNLVGSAASPKCISTEQALLTFRALTDLSSR-----LFSWSGPDCCNWP 68

QY 67 VGCETASGRVVA--LRLP-----KRG--LGGI----- 89  
DB 69 VLCDARTSHVVKIDLRNFSQDVRSDEYKRGSLRGKIHPSLTQLKFLSYLDLSSNDFNEUE 128

QY 90 IPSSIGELDHLRYLDLSGNLSLVEVPKSL-----QIR- 121  
DB 129 IPEFIGQIVSLRYLNLSSSFSGEIPTSLGNLSKLESLDLYAESFGDSGTLSLRASNLRW 188

QY 122 LKSLTTDSQSLSGMSINML-----LHVSSRRTLDEEPNTISG----- 158  
DB 189 LSSLSSTLKYLMGVNLSGAGETWQDFSRISALKELHFNSE-LKNLPPTLSSSADLK 247

QY 159 -----TNNVSGSNVNVSGNDN-----TVVSGNNN-----HVSNNNTV 193  
DB 248 LLEVLDSLNSLNSIPNWLFGLTNLRKLFWRWDFLQSGIPTGFKNLKLETLDSLNNLA 307

QY 194 VTGSNTVVG-----SNHVSGTYKHIVTDNNNVVSGN-----DNNVSGSFHT 235  
DB 308 LQGEIPSVLGDLPQLKFLDLSANELNGQIHGFLDAFSRNKGNLVLFDLSSNKLGLTPE 367

QY 236 VSG---EHTVSGSNNTVSGS 253  
DB 368 SLGSLRNLTDLSSNSFTGS 388

Search completed: October 8, 2005, 02:20:26  
Job time : 43 secs

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: H84421

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-716 <STO>

A;Cross-references: UNIPROT:Q9ZU46; GB:AE002093; NID:g4262228; PIDN:AAD14521.1; GSPDB:GN

C;Genetics:

A;Gene: At2g01210

A;Map position: 2

Query Match 11.9%; Score 164.5; DB 2; Length 716;

Best Local Similarity 27.4%; Pred. No. 0.00017;

Matches 63; Conservative 34; Mismatches 78; Indels 55; Gaps 9;

QY 10 MLAVFLGFTLQVAGATSWSCHHDDHALRGLAENLSGKGAVRLRAAW--SGASCCSWEGV 67

DB 1 MLASLIFFVALLCNVTVISGLNDEGFALLTFKQSVHDDPTGSLN--WNNSDENACSWNGV 59

QY 68 GCTASGRVVALRPLKRGGLGIIPISSIGELDLHLYLDLSGNSLVGEVP-----KSL 118

DB 60 TKCEL--RVVSLIPRKNLYGSLPSSLGFLSSURLNLSNRFPYGSLSPLQLPHLQGLQSL 117

QY 119 QI-----RLKSLTT--DSQSLGMSINM--LHVSSRRTLDEEPTISGT-- 159

DB 118 VLYGNSFDGSLSEIEGKLKLOTLQDLSONLFGNSLPLSLIQCNRKLTLDVSRNLSGPLP 177

QY 160 -----NNSVSGSNVNVSGNDNTVTVSGNNHVS 189

DB 178 DFGSAFVSLKLDLAFNPQNGSIPSDIGNL--SNLQGTADFSHNHFTGS 225

RESULT 12

B85440

receptor kinase-like protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Aug-2004

C;Accession: B85440

R;Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: B85440

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-766 <STO>

A;Cross-references: UNIPROT:O23161; GB:NC\_001268; NID:g7270708; PIDN:CAB80391.1; GSPDB:G

C;Genetics:

A;Gene: AT4g37250

A;Map position: 4

C;Superfamily: protein kinase homology

Query Match 11.8%; Score 164; DB 2; Length 766;

Best Local Similarity 28.5%; Pred. No. 0.0002;

Matches 63; Conservative 35; Mismatches 67; Indels 56; Gaps 9;

QY 60 SCCSWGVCETASGRVVALRPLKRGGLGIIPISSIGELDLHLYLDLSGNSLVGEVPKSL- 118

DB 51 SPCSWRGISCNNDN--KVLTLSPNSQLGSLIPDSGLSLTLQSLDLNNSFNGPLPVSPF 109

QY 119 ---QIRKSLITDSQS----LGMGSINMLHVSRRTLDEEPTISGTNNSVSGSNVV 171

DB 110 NARELFLDSSNNWISGEIPSAIGDLHNL-----TLNLSDNALA-----GKLPNTLA 157

QY 172 SGNDNTVSGNNHVSNNVTVGSDNTVVGSHVSGTKHIVTNNNVVSG----- 224

DB 158 SLRNLTVVSLNNYFSGE-----IPGWRVV---EFLDLSNLSLNGSLPPDFG 202

QY 225 --NDNVSGSFHTVSGE-----HNTVSGSNNTVSG 252

DB 203 GYSLOYLVNSFQISGEIPPEIGVNPFRNVTVDLSFNLTG 243

RESULT 13

T05322

hypothetical protein F18F4.240 - Arabidopsis thaliana

N;Alternate names: hypothetical protein F1C12.60

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T05322; T04898

R;Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, April 1998

A;Reference number: Z15408

A;Accession: T05322

A;Molecule type: DNA

A;Residues: 1-1232 <BEV>

A;Cross-references: UNIPROT:Q9SN91; EMBL:AL022224

A;Experimental source: cultivar Columbia; BAC clone F1C12

R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, February 1998

A;Reference number: Z15388

A;Accession: T04898

A;Molecule type: DNA

A;Residues: 1-305 <BEW>

A;Cross-references: EMBL:AL021637

A;Experimental source: cultivar Columbia; BAC clone F18F4

C;Genetics:

A;Map position: 4

A;Intons: 863/1; 1116/1

A;Note: F1C12.60; F18F4.240

C;Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology;

Query Match 11.7%; Score 162.5; DB 2; Length 1232;

Best Local Similarity 25.3%; Pred. No. 0.00045;

Matches 71; Conservative 41; Mismatches 94; Indels 75; Gaps 11;

QY 11 LLVFLGFTLQVAGATSWSCHHDDHAL-----RGLAENLSGKGAVRLRAAWS--GASCCSW 64

DB 4 LVLLLLFLLFCFGLGQGIINNDLQTLLEVKKSLVTNPQEDDPLR---QWNSNINVCWS 60

QY 65 EGVGCB--TASGRVVALRPLKRGGLGIIPISSIGELDLHLYLDLSGNSLVGEVPKSLQIRLK 123

DB 61 TGYTCNTGLFRVIALNLTLGLTGSISPFWGFEDNLHLDLSSNLVGPITALS-NLT 119

QY 124 SLTTDSQSLGMSINMLHVSRRTLDEEPTISGTNN--SVGSGSNVNVSGNDNTVVG 181

DB 120 SLES-----LFLFSNQLTGEIPSQLGSLVNIRSLRIGDNELVGDIPETL--G 164

QY 182 N-----NNHVS-----SNTVTGSD 198

DB 165 NLVNLQMLALASCLTGTPIPSQLGRVVSQSLIQDNLYLEGPIPAELGNCSDLTVTFAAE 224

QY 199 NTVVSGSNVVSQTKHIVTNNNVVSGNDNVVSGSFHTVSGE 239

DB 225 NMLNGTIPAEGR-----LENLEILNANSLTGEIPSQLGE 261

RESULT 14

T00502

probable receptor-like protein kinase At2g23300 [imported] - Arabidopsis thaliana

N;Alternate names: protein kinase homolog T2016.7

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 16-Aug-2004

C;Accession: T00502; A84623

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau

submitted to the EMBL Data Library, November 1997

A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.

A;Reference number: Z14164

A;Accession: T00502

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-773 <ROU>

C; Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C; Accession: AE1857  
R; Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A; Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anaerostipes*  
A; Reference number: AB1807; MUID:21595285; PMID:11759840  
A; Accession: AE1857  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-420 <KUR>  
A; Cross-references: UNIPROT:Q9YZQ1; GB:BA000019; PIDN:BAB72364.1; PID:G17129751; GSPDB:G  
A; Experimental source: strain PCC 7120  
C; Genetics:  
A; Gene: all0406

Query Match	12.2%	Score 169;	DB 2;	Length 420;
Best Local Similarity	31.5%;	Pred. No. 4e-05;		
Matches	Conservative 28;	Mismatches 51;	Indels 34;	Gaps 8
127	TDSQSLGWSINMLLHVSSRRLLDEEPTISCTNN-SVGS-----GSSNVVSGN----	174		
130	SDNATIGNGWNF-----NTNNTTVNGWNLFGSSNTTIGNGWYWDGSSNATLGNGNWH	185		
175	---DNTVV-----SGNNHHVSGSNNTTVTGSDNTVVGSHVHVSGTGHVTDNNNVVS	223		
186	FGSDNRTTICNGWYLDGNNNATLGNGNWHY-GSDNTTIG-----NGWYLDGNNNATL	239		
224	GNDDNVSGSFHTVSGEHTVSGSNNTVSGS-NHIVSGSNKVVTDG	267		
240	GNGNWYFGTDNTTICNGWWDGFTNNITLGNGNWIFETNNTVVGN	284		

RESULT 9  
G96746  
hypothetical protein T9N14.20 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: G96746  
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizlar, L.  
nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: G96746  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1095 <STO>  
A;Cross-references: UNIPROT:Q9C7S5; GB:AE005173; NID:gl0645385; PIDN:AAG21504.1; GSPDB:G  
C;Genetics:  
A;Gene: T9N14.20  
A;Map position: 1

	Query Match	12.2%;	Score 168.5;	DB 2;	Length 1095;
	Best Local Similarity	28.7%;	Pred. No. 0.00014;		
	Matches	58;	Conservative 29;	Mismatches 74;	Indels 41; Gaps 9;
QY	17	FILQVAGATSWSCHDDLHALRGALNLSKGA VRLRAAM--SGASCCSWGVGCE--TASG	74		
		:   :   :   :   :			
Db	39	FFITVSEAV--CNLQDRDSLLWFGSNVSPVS---PLHNSSITDCCSWGEGISCDKSPEN	92		
		:   :   :   :   :			
QY	75	RVALRPLRKGLGIIPSSIGELDHLYRLDLGNSLVGEVPKSLQLRIKLSTLTDSQS LGM	134		
		:       :       :       :			
Db	93	RVTSIILSRGLSGNLPPSVLDLQLRLDLSHNRLSGPLPGF-----L	137		
		:       :       :       :			
QY	135	GIMMLHVVSSRRTLPDEPNTISG---TNNSVSGSGNNV-----VSGN--DNMTVVSGN	182		
		: : :       :       :       :			

Db 138 SAlDQLL-----VLDSLNSFKGELPLQQSGFGNGSIPIFIQTVDLSSNLEGEILSSS 191

Qy 183 NNHVGSGNNVTVTGSDNVTVGS 204

Db 192 VFLOGAFLNLTFSNVSNNSFTGS 213

RESULT 10

C96519

probable disease resistance protein, 1954-7013 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: C96519

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; De-

ansen, N.F.; Hughes, B.; Huiziar, L.

Nature 408, 8167-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C96519

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1019 <STO>

A;Cross-references: UNIPROT:Q9C699; GB:AE005173; NID:G10092530; PIDN:ANG12927.1;

C;Genetics:

A;Gene: TGB12.2

A;Map position: 1

Query Match	12.1%;	Score 168;	DB 2;	Length 1019;
Best Local Similarity	23.2%;	Pred. No. 0.00014;		
Matches	84;	Conservative 50;	Mismatches 106;	Indels 122; Gaps 14;
Qy	9	CMLVFLGFILOVAGATSWCSHDDHIALRGLAENLSKGAVRLRAAWSGASCCSWEGVG	68	
Db	54	CFILIPSLFIITVFSATQHLCHSDQKDALLDFKNEF---	GWVDSKSWYKSDCCSWDG	110
Qy	69	CETASGRVVVALRLPKRGLGGIIP--SSIGELDHLR-----		101
Db	111	COAKSGNVIGLDSSTIFYGQKSNSSLFKLRLRDLNANNFNNSPIPAEFDKLTGLE	170	
Qy	102	YLDLSCNSIAGVEPKSLQIRLKSLLTDSQ-----		131
Db	171	RDLDSOSSLSGGIPINLQLTKVLSDLSDSSDFGDESFHYLSIDKSFPLRLARNLRLR	230	
Qy	132	-LQMGSGINN-----LLHVSSRRITD-----EEPNTISGTNN--SVGSGSNVVVSG	173	
Db	231	ELDMSYVKISSEIPEEFSNIRSLRSLNLCNLGCFGEFFSSILLIPNLQSIDLGNPNLRG	290	
Qy	174	NDMTVVSGNNHV-----SGS-----NNTVVGSDNTTVVG-----SNHVVSG	210	
Db	291	--NLVPFHENNSLLKILTYTFSGAIPDSISSLKNLTSLTSLSVSYFGKIPFSLGNLUSH	348	
Qy	211	TKHIVTDNNWVS-----GDNNNVSGSP---HTVSG-----BHNTVSGSNNTVS	251	
Db	349	LSHLSSNNLICEIPSSIGNLNQLTNFVVGNGKLSGNLPATLSNLTKLNTISLSNQFT	408	
Qy	252	GS	253	
Db	409	GS	410	

RESULT 11

H84421  
probable receptor-like protein kinase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: H84421  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii

C;Accession: G84524  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: G84524  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-983 <STO>  
A;Cross-references: UNIPROT:Q9ZUK3; GB:AE002093; NID:g4115363; PIDN:AAD03365.1; GSPDB:GN  
C;Genetics:  
A;Gene: At2g15080  
A;Map position: 2

Query Match 13.0%; Score 179.5; DB 2; Length 983;  
Best Local Similarity 21.7%; Pred. No. 1.9e-05;  
Matches 81; Conservative 56; Mismatches 113; Indels 123; Gaps 14;

QY 1 MPEYMAKCCMLLVFLGFILOVAGATSWSCHHDDLHAL-----RGLAENLSGKGAVALR 53  
DB 2 MKGYITLSFLIILIFNLFDEFAASTRHLCDDPDQSDAILFEKNEFETLEBSCFDSNIPKLT 61

QY 54 AAWSGAS-CCSWBEGVGCETASGRVVAL-----RPL----- 82  
DB 62 ESWTNSDCYWDGIKDAKFGDVIELDSFSCRLGQLNSNSLFLRLPOLRFLTTLDLSN 121

QY 83 -----KRGLGIIPISSIGELDHLRLDLSGNSLVGEVPKSIQ 119  
DB 122 NDFIQIPISSLETLSNLTLDLSRNHFSGRIPISSIGNLSHLIFVDFSHNFSQIPISSLG 181

QY 120 IRLKSLTTDSQSL-----GMGSINML--LHVSRRTLDEEPTISG----- 158  
DB 182 Y-LSHLTSPNLSNPNFSGRPSSIGNLSYLTTLRLSRNFFGGLPSGLSFLHLTDLILD 240

QY 159 TNNVSGSGNNVSGNDNTVSGNNHNSGNNVTGSDNTVSGNHVVS-----TKH 213  
DB 241 THNFVKIPSSL-----GNLSHL-----TSDLHKNNFVGPIPSGLNSCLTGF 285

QY 214 IVTDNN-----NVSGNDNNVSGSFHTV---SGEHTVSGNNVTSGS--- 253  
DB 286 ILSDDNIVGEIPSSFGNLAQLDILNVKNSKLSGSPFIALLNLRLKLTSLSLFNRLGTLP 345

QY 254 NHIVSGSNKVVD 266  
DB 346 SNMSLSNLKLF 358

RESULT 6  
C84633  
probable disease resistance protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: C84633  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: C84633  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-743 <STO>  
A;Cross-references: UNIPROT:Q9ZUH7; GB:AE002093; NID:g4115376; PIDN:AAD03377.1; GSPDB:GN  
C;Genetics:  
A;Gene: At2g24160  
A;Map position: 2

Query Match 12.7%; Score 176.5; DB 2; Length 743;  
Best Local Similarity 26.7%; Pred. No. 2.3e-05;  
Matches 63; Conservative 35; Mismatches 81; Indels 57; Gaps 10;

QY 55 AAWSGAS-CCSWBEGVGCETASGRVVALR-----PKRG-----LG 87  
DB 11 SNKSSDCCSWBESVTCDAKYGQVISLLGVNLNNTLKNPSGLFKLQYLQNLTLRYCNLY 70

QY 88 GIPISSIGELDHLRLDLSGNSLVGEVPKS-----LQIRLKSLLTTDSQSLGMSIN 138  
DB 71 GEIPFSLGTLSHLTFLDLSENKLVGVQPSISIGNLTKLMYLRLSINHLSGKS-SVSPANLT 129

QY 139 MLLHVSSRRTLDEEPTI---SGTNNVSGSGNNVSGNDNTVVS-----GNNN 184  
DB 130 KLIQIDIREN-DFEPELIPDMRFRHNLGFGGNGFFGPTSLFTIPSLRWVNLDRSNNN 188

QY 185 ---HVSGSNNVTGSDNTVSGNHVSGTKHIVTDNNNV--SGMNDNNVSGSFHT 235  
DB 189 FTGHIDFGNSSLSRSLSYLSLADNPFDPPIPEISIKFLNLVLLDLRNSFSGFPT 244

RESULT 7  
B86440  
probable protein kinase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: B86440  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.P.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: B86440  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-590 <STO>  
A;Cross-references: UNIPROT:Q9C867; GB:AE005172; NID:g11054670; PIDN:AAG27890.1; GSPDB:  
C;Genetics:  
A;Map position: 1

Query Match 12.5%; Score 173.5; DB 2; Length 590;  
Best Local Similarity 24.7%; Pred. No. 2.8e-05;  
Matches 70; Conservative 44; Mismatches 90; Indels 79; Gaps 13;

QY 5 MAKCCMLLVFLGFILOVAGATSWSCHHDDLHALRGLAENLSGKGAVALRAA----- 55  
DB 6 MKSCCSWLLLSILCSLSNESQ-----AISPDGEALLSFRNAVTRSDSFH 51

QY 56 -W-SGASCCSWBEGVGCETASGRVVALRIPKRGLGIIPISSIGELDHLRLDLSGNSLVG 112  
DB 52 QWRPEDPDCCNNGVTCDAKTRVITLNTYHKIMGPLPDPDICKDLHLRLMLHNNALYG 111

QY 113 EYPKSL-----QIRLKS-LTTDSQSLGMSINMLLVSSRRTLDEEPTISG----- 158  
DB 112 AIPTALGNCTALEEHLQSNYFTGPIPAEMGDPLG-----OKLMSNLTSLGPIPASL 165

QY 159 -----TNNVSGSGNNVSGN--DNTVSGNNHNSGNNVTGSDNTVSGNHVVS 210  
DB 166 GQLKLSNFNV---SNNFLVGQIPSDGVLSG-----FSKNSFIGNLNLG- 207

QY 211 TKHIVTDNNNVSGNDNNVSGSFHTVSGEHTVSGNNVTSGS 253  
DB 208 -KHV-----DVVCQDDSG-NPSSHSGSQGNQKNSGKLLISAS 243

RESULT 8  
AE1857  
hypothetical protein al10406 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

R.Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;   
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,   
 Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AF0169

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-622 <KUR>

A:CROSS-references: UNIPROT:O8ZGB5; GB:AL590842; PIDN:CAC90217.1; PID:g15979437; GSPDB:G

C:Genetics:

A:Gene: YPO1388

Query Match 13.7%; Score 190; DB 2; Length 622;  
 Best Local Similarity 26.3%; Pred. No. 1.8e-06;  
 Matches 68; Conservative 49; Mismatches 72; Indels 70; Gaps 13;

QY 43 NLGSGAVLRAAWGASCCSWEVGCETASGRVVALRPKGLGLIIPSSIGELDLHRY 102

DB 41 NNGNGTINIFDASSNNDIHTLTGLNEQ-----LGGF-----SNHL-- 77

QY 103 LLDLGSNLVGEVPSLQIRLKSITDSQ-----SLQMG-----SINMLHVSS 145

DB 78 IDSHNTIDGGQSNL-----VSDGNMISALSLDGLFYGAQNTTLNSNNLLIVTQ 131

QY 146 RRT-LDEEPTISGTNNVSGSNNVSGNDTVVS-----GNNHVSGSNTV 193

DB 132 GSTIIDSNTVSGISNNLIESNNII-GNENSCYSDPASPGAWCVDNQLTIGSDNNT 190

QY 194 VTGSDNTVGSNH--VVSSTKHIVTDNNVSGNDNNVSGSFHTVS-CEHNTVSGSNT 249

DB 191 ITGALNGLNHHNDIIASVNNMLDTHNIIAG-----GHNTISGGNNDIFGSENN 244

QY 250 VSGSNHI-VSGSNKVYTDG 267

DB 245 VTDSIDANINGSNNYIDG 263

RESULT 3

T05251

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T05251

R:Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; Maye

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15405

A:Accession: T05251

A:Molecule type: DNA

A:Residues: 1-645 <BEV>

A:CROSS-references: UNIPROT:O9SVN1; EMBL:AL035528

A:Experimental source: cultivar Columbia; BAC clone F18A5

C:Genetics:

A:Map position: 4

A:Introns: 216/1; 541/1

A:Note: F18A5.210

Query Match 13.4%; Score 186; DB 2; Length 645;

Best Local Similarity 22.1%; Pred. No. 3.8e-06;

Matches 82; Conservative 46; Mismatches 115; Indels 128; Gaps 14;

QY 9 CMLVFLGFLQVAGATSWSCCHDDHLALRGLAE-----NLGSGAVRLRAAW-SGAS 60

DB 7 CLILSLNSKLVLASHVGHLCRQDKNALLEFKNFEVHFENSGIVGVKTEKWRNNTD 66

QY 61 CCSWEGVGCETASGRVVALRL-----PKR-----GLGIIIPSSI 94

DB 67 CCSWDGISCDPKTKGVVELDLNMSFLNGPLRYDSSLFRLOHLNLDLGSNNFSGILPDSI 126

QY 95 GELDHLRYLDLGSNLVGEVPSLQIRLKSITDSQSL-----GMGSINML--LHV 143

DB 127 GSKYLRLVSLGDCNLFKGIIPSSILG-NLTYLTNLDLSVNDFTGELPDSMGHLNKLTELHL 185

QY 144 SSRTLDDEEPTISGTN--NSVSGSN-----NVVSGN-----DN 176

DB 186 GSAKLSGNFPMLNLSELTLIDSGNOFGENQTYLDISANKIGQVQPWMLSLPEL 245

QY 177 TVVSGNNHVSQ-----SNNTVVTGSDNTVVS 204

DB 246 QYVNISQNSFSFGFEGPADVIQRCBELLMLDISSNTFFQDPFLLPNSTIFLGSNDNRFSGE 305

QY 205 -----NHVVSGTKHIVTDNN-----NVVSGNDNNVSGSF--HTVSGEHN 241

DB 306 IPKTIKLVSLDTLVLSNNFNNGSIPRCFEKFNNTLSVLHRLNNLSGEFPESISDHUR 365

QY 242 TVSGSNNTVSG 252

DB 366 SLDVGRNRLSG 376

RESULT 4

T46070

hypothetical protein T18N14.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 16-Aug-2004

C:Accession: T46070

R:Delser, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z23013

A:Accession: T46070

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-836 <DEL>

A:CROSS-references: UNIPROT:O9SCT4; EMBL:AL132968

A:Experimental source: cultivar Columbia; BAC clone T18N14

C:Genetics:

A:Map position: 3

A:Introns: 625/1

A:Note: T18N14.120

C:Superfamily: protein kinase homology

Query Match 13.3%; Score 184.5; DB 2; Length 836;

Best Local Similarity 25.7%; Pred. No. 6.7e-06;

Matches 82; Conservative 48; Mismatches 100; Indels 89; Gaps 15;

QY 10 MLLVFLGFLQV-----AGATSW-----SCHDDHLALRGLAENLSGKAVRLRAWSG 58

DB 22 LLLCFLFSAQAVAGGGGSHDGIIVTQYQALQAIKHELIDFTG-----VLKSWNN 76

QY 59 AS-----CCSWEGVGCETASGRVVALRPKRGIGGIIPSSIGELDLHRYLDLGSNLVGEV 114

DB 77 SASSQVCSGAGIKC--LRGQVVAIQLPWKGLGGTISEKIGQLGSLKLSLHNNVIAGSV 134

QY 115 PKSLQIRLKS-----TTDSQSLGMSINMLLVSSRRTLDEEPTISGT----- 159

DB 135 PRSLGY-LKSLRGVYLFNNRLSGSIPVSLGNCPLL-----QNLDSLNLQLTGAIPT 187

QY 160 -----NNSVSGSNNVSGNDNTVVSNNNHVSGS--NNTVVTGS-----D 198

DB 188 ESTRLRYLNLSPNSLSGPLPVSVARSYTLTFLQHNLSGIPDFVNGSHPLKTLNLD 247

QY 199 NTVVSGSNHVSSTKHIVTDNNNNVSGNDNNVSGSFHTVSG-----EHNVTSGS-- 246

DB 248 HNRFGSAPVSLCKHSLLEE---VSISHNLQSSIPRECGLPLHQSLDYSYNGITIP 304

QY 247 -----NNTVS---GSNHI 256

DB 305 DSFSNLSLVLNLSNHL 323

RESULT 5

G84524

probable disease resistance protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2005, 02:09:36 ; Search time 41 Seconds  
(without alignments)  
626.582 Million cell updates/sec

Title: US-10-657-852B-15  
Perfect score: 1385  
Sequence: 1 MPEYMAKCMCLLVFLGLIQ.....NTVSGSNHIVSGSNKVVTDG 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: Pirl:\*  
2: Pirl2:\*  
3: Pirl3:\*  
4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	190.5	13.8	1008	2 D84434	probable receptor-
2	190	13.7	622	2 AF0169	probable exported
3	186	13.4	645	2 T05251	probable disease r
4	184.5	13.3	836	2 T46070	hypothetical prote
5	179.5	13.0	983	2 G84524	probable disease r
6	176.5	12.7	743	2 C84633	probable disease r
7	173.5	12.5	590	2 B86440	probable protein k
8	169	12.2	420	2 AE1857	hypothetical prote
9	168.5	12.2	1095	2 G96746	hypothetical prote
10	168	12.1	1019	2 C96519	probable disease r
11	164.5	11.9	716	2 H84421	probable receptor-
12	164	11.8	766	2 B85440	receptor kinase-li
13	162.5	11.7	1232	2 T05322	hypothetical prote
14	162	11.7	773	2 T00502	probable receptor-
15	161.5	11.7	905	2 T00475	probable disease r
16	161	11.6	719	2 T47727	hypothetical prote
17	161	11.6	907	2 A86460	99.9K hypothetical
18	159.5	11.5	768	2 T17462	disease resistance
19	158.5	11.4	658	2 B84664	probable receptor-
20	158	11.4	478	2 H86459	hypothetical prote
21	157.5	11.4	996	2 T10725	protein kinase Xa2
22	157	11.3	241	2 S61925	cyst wall protein
23	156.5	11.3	702	2 A86383	76.4K protein kin
24	156.5	11.3	876	2 D96558	probable protein k
25	156	11.3	371	2 T49908	hypothetical prote
26	156	11.2	638	2 T05606	protein kinase hom
27	155.5	11.2	942	1 JQ1674	protein kinase TMK
28	155	11.2	729	2 F86308	Similar to disease
29	154.5	11.2	1143	2 T10636	hypothetical prote

30	154	11.1	720	2 T02361	hypothetical prote
31	153	11.0	668	2 T05257	probable disease r
32	152.5	11.0	646	2 B84852	probable receptor-
33	152.5	11.0	655	2 G96524	protein Tln15.9 [i
34	152.5	11.0	838	2 A96557	probable receptor-
35	152.5	11.0	980	2 H84632	probable disease r
36	151.5	10.9	480	2 T00971	probable disease r
37	151	10.9	693	2 F96557	probable protein k
38	151	10.9	863	2 F85343	hypothetical prote
39	151	10.9	981	2 T50851	receptor protein k
40	150.5	10.9	221	2 T07079	leucine-rich repea
41	150.5	10.9	672	2 B84782	probable receptor-
42	150.5	10.9	813	2 T04313	protein kinase Xa2
43	149.5	10.8	286	2 C61615	sericin MG-2 - gre
44	149.5	10.8	645	2 E96631	probable receptor-
45	149	10.8	976	2 B84659	probable receptor-

ALIGNMENTS

RESULT 1

D84434  
probable receptor-like protein kinase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: D84434  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.C.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10611797  
A:Status: preliminary  
A:Accession: D84434  
A:Molecule type: DNA  
A:Residues: 1-1008 <STO>  
A:Cross-references: GB:AE002093; NID:96598459; PIDN:AAC78507.2; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g02220  
A:Map position: 2

Query Match	13.8%;	Score 190.5;	DB 2;	Length 1008;
Best Local Similarity	27.5%;	Pred. No. 3e-06;		
Matches	68;	Conservative	45;	Mismatches 93;
Indels	41;	Gaps	12;	
QY	9	CMLLVFLG----	FILQVAGATSWSCHDDDLHALRGLAENLSGKGAVRLRAAWSGASCCSW	64
Db	7	CVIVIFLTLLCFFYSSQTTSRCHPHDLEALRDFIAHLEPKPDGWINSS--SSTDCNW	65	
QY	65	EGVGCT-ASGRVVALRPLKRGIGGIIPSSIGELDHLRYLDLSGNSLVGEVPKSLQIRLK	123	
Db	66	TGTCNNNTGVRILELGNKKLSGLKSLGKLDLRVLNLSRNFKDSIPLSI--FNLK	124	
QY	124	SILT-----DSQSLGMSINLLHVSSRRTLDEPNTISGT-----NNS-----VGSGSNV	170	
Db	125	NIQLTDLSSNDLS--GGIPTSINLPALQSFDSLNSKFNGLPSHICHNSTQIRVVKLAVNY	183	
QY	171	VSGNDNTVVSNNHVSNNVTVTGSDNTVVSNNHVSSTKHIVTD-----NNNVVSGN	225	
Db	184	FAG-----NFTSGFGKCVLL---EHLCLGMNDL---TGNIPEDLFLKELNLLGIQ	228	
QY	226	DNVSGS 232		
Db	229	ENRLSGS 235		

RESULT 2

AF0169  
probable exported protein YP01388 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AF0169

**This Page Blank (uspto)**



PD 18-MAR-2004.  
 XX  
 PF 09-SEP-2003; 2003WO-NZ000199.  
 XX  
 PR 09-SEP-2002; 2002US-0409557P.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX (WRIG-) WRIGHTSON SEEDS LTD.  
 PA  
 XX Demmer J, Shenk MA, Hall C, Fish SA;  
 PI WPI; 2004-248453/23.  
 XX N-PSDB; ADM41458.  
 DR  
 XX New antifreeze proteins and encoding polynucleotides, useful for  
 PT modulating cold tolerance in organisms, as food additives, or for  
 PT treating tumors or disorders associated with the presence of unwanted  
 PT biocrystals (e.g. gout).  
 XX  
 XX Claim 7; SEQ ID NO 14; 71pp; English.  
 PS  
 XX The present sequence is that of APF1, an antifreeze protein of perennial  
 CC ryegrass. The encoding cDNA was isolated from a pseudostem cDNA  
 CC expression library. The invention provides forage grass (perennial  
 CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides  
 CC encoding them ADM41458-ADM41483. The polynucleotides were isolated from  
 CC tissues taken at different times of year (winter and spring) and from  
 CC different parts of the plants. The polynucleotides can be used to  
 CC modulate the cold tolerance of an organism, especially plants, mammals,  
 CC insects, fungi, archaea and bacteria. The method involves incorporating  
 CC an antifreeze polynucleotide, under the control of a gene promoter  
 CC sequence, into the genome of the organism, or introducing double-stranded  
 CC RNA corresponding to the polynucleotide into the cells of the organism,  
 CC thereby inhibiting expression of an antifreeze polypeptide. The  
 CC antifreeze protein can be used for the cryopreservation of a cell or  
 CC tissue, as a food additive of a frozen food product, in a method for  
 CC decreasing the time required to dehydrate a composition, to treat a  
 CC disorder characterised by biocrystals associated with disorders such as  
 CC gout and kidney stones, to preserve the viability of a molecular biology  
 CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
 CC to protect a plant from damage due to frost or freezing.  
 XX  
 XX Sequence 243 AA;  
 Query Match 47.3%; Score 655; DB 8; Length 243;  
 Best Local Similarity 51.9%; Pred. No. 1.8e-49;  
 Matches 148; Conservative 18; Mismatches 55; Indels 64; Gaps 5;  
 QY 5 MAKCCMLVFLGLIQLVAGATSWCHDDHLHALRGLAENLSGKGVRLRAAWSGACSCSW 64  
 DB 1 MAKCWQLLLFLALLPAASAA--SCHPDDLYALRDFAGNLRG--GGVLLRAALPGASCCGW 57  
 QY 65 EGVGCETASGRVVALRPLKRGGLGIIIPSSIGELDLHLYLDLSGLNLSVGEVPSKLRILKS 124  
 DB 58 EGVGCDGASGCV-----KSFQILLKG 78  
 QY 125 LTTDSQSLGMSINMLLHVS--SRRLDEEPNTISGTTNNSVGSNNVSGNDNTVVSNN 183  
 DB 79 LTAAGRSGLKAFTHPLHVRPSQGLTDEHNTTIGINTVRSNNVSGNDNTVVSNN 138  
 QY 184 NHVSGSNNTVVTGSDNTVGSNNHVSGTKHIVTDNNNVSGNDNNVSGSPHTVSGEHTV 243  
 DB 139 NVVSGSHNTVVFQDGNFISGYHVSGNHHVVTDNKNVSGDNTVSGSQNTVSGNHQIV 198  
 QY 244 SGS-----NNTVSGSNHIVSGSNKVVTDG 267  
 DB 199 SGSHSTVSGNHNNTVSGRNNVSGNNIVSGSNHVGVGNKVVTDG 243  
 RESULT 14  
 AAY22472  
 ID AAY22472 standard; protein; 118 AA.  
 XX

AC AAY22472;  
 XX 29-SEP-1999 (first entry)  
 XX  
 DE Grass anti-freeze protein sequence.  
 XX  
 KW Anti-freeze protein; grass; plant; frozen food product; frost tolerance;  
 KW frozen confectionery.  
 XX  
 OS Lolium perenne.  
 XX  
 PN WO9937782-A2.  
 XX  
 PD 29-JUL-1999.  
 XX  
 PF 23-DEC-1998; 98WO-EP008553.  
 XX  
 PR 22-JAN-1998; 98GB-00001408.  
 XX  
 PA (UNIL ) UNILEVER NV.  
 PA (UNIL ) UNILEVER PLC.  
 XX  
 XX Jarman CD, Sidebottom CM, Twigg S, Worral D;  
 PI WPI; 1999-458697/38.  
 XX N-PSDB; AAX99717.  
 DR  
 XX New plant anti-freeze protein useful in frozen food products.  
 PT  
 PS Claim 3; Page 36; 39pp; English.  
 CC  
 CC This sequence is the plant anti-freeze protein of the invention. The anti  
 CC -freeze protein is characterised in that at least 40% of its amino acids  
 CC are from the group of serine, threonine and asparagine. The anti-freeze  
 CC protein can be used in frozen food products, especially frozen  
 CC confectionery. Anti-freeze proteins are especially used in food products,  
 CC which are heated, e.g. by pasteurisation, blanching or sterilisation  
 CC prior to freezing. Plants transformed with a nucleic acid sequence  
 CC encoding the anti-freeze protein have an increased frost tolerance. Prior  
 CC art anti-freeze proteins have not been applied to commercially available  
 CC food products, due to high costs and complicated processes for obtaining  
 CC the protein. Also prior art anti-freeze proteins have tended to  
 CC destabilise during processing especially during the pasteurisation step.  
 CC This is overcome by the present anti-freeze protein. The anti-freeze  
 CC proteins provide an ice particle size following an ice recrystallisation  
 CC inhibition assay of 15 µm M or less. The anti-freeze protein ingredient  
 CC means that mixes can be frozen under quiescent conditions, e.g. in a shop  
 CC or home freezer without the formation of unacceptable ice crystal shapes  
 CC and hence with a texture different to products normally obtained via  
 CC quiescent freezing  
 XX  
 XX Sequence 118 AA;  
 Query Match 38.3%; Score 530; DB 2; Length 118;  
 Best Local Similarity 85.5%; Pred. No. 7.3e-39;  
 Matches 100; Conservative 10; Mismatches 7; Indels 0; Gaps 0;  
 QY 150 DREPNTISGTTNNSVGSNNVSGNDNTVVSNNHVSNNNTVVSNDNTVVSNNHVS 209  
 DB 1 DEQPNTISGNNTVRSKXVLAGNDNTVSGDNNVSGNNNTVVSNDNTVVSNNHVS 60  
 QY 210 GTKHIIVTDNNNVSGNDNNVSGSFHTVSGEHTVSGNNHVSNNHVSNNHVSNNHVS 266  
 DB 61 GTNHIIVTDNNNVSGNDNNVSGSFHTVSGEHTVSGNNHVSNNHVSNNHVSNNHVS 117  
 RESULT 15  
 ABB93804  
 ID ABB93804 standard; protein; 1036 AA.  
 XX  
 AC ABB93804;  
 XX  
 XX 31-MAY-2002 (first entry)  
 DT



PR 09-SEP-2002; 2002US-0409557P.  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
PA (WRIG-) WRIGHTSON SEEDS LTD.  
PI Demmer J, Shenk MA, Hall C, Fish SA;  
XX WPI; 2004-248453/23.  
DR N-PSDB; ADM41466.  
XX  
PT New antifreeze proteins and encoding polynucleotides, useful for  
PT modulating cold tolerance in organisms, as food additives, or for  
PT treating tumors or disorders associated with the presence of unwanted  
PT biocrystals (e.g. gout).  
XX  
PS Claim 7; SEQ ID NO 22; 71pp; English.  
XX  
CC The present sequence is that of an antifreeze protein of tall fescue. The  
CC encoding cDNA was isolated from a leaf blade cDNA expression library. The  
CC invention provides forage grass (perennial ryegrass and tall fescue) antifreeze  
CC proteins and the polynucleotides encoding them ADM41458-  
CC ADM41483. The polynucleotides were isolated from tissues taken at  
CC different times of year (winter and spring) and from different parts of  
CC the plants. The polynucleotides can be used to modulate the cold  
CC tolerance of an organism, especially plants, mammals, insects, fungi,  
CC archaea and bacteria. The method involves incorporating an antifreeze  
CC polynucleotide, under the control of a gene promoter sequence, into the  
CC genome of the organism, or introducing double-stranded RNA corresponding  
CC to the polynucleotide into the cells of the organism, thereby inhibiting  
CC expression of an antifreeze polypeptide. The antifreeze protein can be  
CC used for the cryopreservation of a cell or tissue, as a food additive of  
CC a frozen food product, in a method for decreasing the time required to  
CC dehydrate a composition, to treat a disorder characterised by biocrystals  
CC associated with disorders such as gout and kidney stones, to preserve the  
CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
CC patient e.g. tumour tissue, and to protect a plant from damage due to  
CC frost or freezing.  
XX  
SQ Sequence 285 AA;  
Query Match 65.1%; Score 901.5; DB 8; Length 285;  
Best Local Similarity 65.5%; Pred. No. 3.5e-71;  
Matches 186; Conservative 17; Mismatches 54; Indels 27; Gaps 3;  
QY 10 MLLVFLGFLQVAGATSWCHHDDLHALRGLAENLSGKAVRLRAAWSGSCSWEVGC 69  
Db :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:|  
3 LLLLPLGLFLPAACATSSCHPDLRALRGFAKVG-GGVLLRTAWSGTSCCWEGVGC 61  
QY 70 ETASGRVVALRLPKRGIG-----GIIPSSIGELDLRLYLVD 104  
Db :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:|  
62 NGASGRITTLWLPRLRGAGTTGASLAGLAELESNLANNRLVGTIPSWIGELDLRLYLVD 121  
QY 105 LGSNSLVGEVPKSLQIRLSLTDSQSLGWSINMLLHVS-SRRTLDEEPNTISGTNNVS 163  
Db :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:|  
122 LSHNSLVGELPNRLQIRLKLTTTGHLLGNAFTNMPDLVKRRTLAIQNTISGTNNLV 181  
QY 164 GSGSNVSGNDNTVVSNNHVSNNNTVTGSDNTVVSNNHVSNGTKHIVTDNNVVS 223  
Db :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:|  
182 LSGRNNVSGNDNTVISENNNTVSGSFNTVTGSDNVLATGSHNVSGRSHIVTDNNVS 241  
QY 224 GNDNNVSGFHTVSGEHTVSGSNTVSGSNNHVSNNKVVTDG 267  
Db :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:| :|||:||||:|  
242 GDDNNVSGSFHKVSGSHNTVSGSNNNTVSGRNVHVSNNKIVTGG 285

RESULT 11  
ADM41478  
ID ADM41478 standard; protein; 280 AA.  
XX  
AC ADM41478;  
XX  
DT 03-JUN-2004 (first entry)  
XX

DE Perennial ryegrass antifreeze protein APP4.  
XX  
KW Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
KW antigout; litholytic; nephrotropic; cytostatic.  
XX  
OS Lolium perenne.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..22 /label= Signal\_peptide  
FT Region 12..22 /note= "Conserved lipoprotein membrane attachment site"  
FT Region 52..63 /note= "Conserved Cys-pairs identified N-terminal or  
FT leucine-rich repeats of receptor-like kinases"  
FT Region 68..134 /note= "Conserved leucine-rich repeats"  
FT Domain 134..153 /note= Transmembrane domain  
FT Region 165..279 /note= "7-amino acid sequence repeat region identified in  
FT antifreeze proteins"  
XX WO2004022700-A2.  
XX 18-MAR-2004.  
PD  
XX 09-SEP-2003; 2003WO-NZ000199.  
PF  
XX 09-SEP-2002; 2002US-0409557P.  
PR  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
PA (WRIG-) WRIGHTSON SEEDS LTD.  
XX  
PI Demmer J, Shenk MA, Hall C, Fish SA;  
XX  
DR WPI; 2004-248453/23.  
DR N-PSDB; ADM41465.  
XX  
PT New antifreeze proteins and encoding polynucleotides, useful for  
PT modulating cold tolerance in organisms, as food additives, or for  
PT treating tumors or disorders associated with the presence of unwanted  
PT biocrystals (e.g. gout).  
XX  
PS Claim 7; SEQ ID NO 21; 71pp; English.  
CC  
CC The present sequence is that of APP4, an antifreeze protein of perennial  
CC ryegrass. The encoding cDNA was isolated from a leaf and pseudostem cDNA  
CC expression library. The invention provides forage grass (perennial  
CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides  
CC encoding them ADM41458-ADM41483. The polynucleotides were isolated from  
CC tissues taken at different times of year (winter and spring) and from  
CC different parts of the plants. The polynucleotides can be used to  
CC modulate the cold tolerance of an organism, especially plants, mammals,  
CC insects, fungi, archaea and bacteria. The method involves incorporating  
CC an antifreeze polynucleotide, under the control of a gene promoter  
CC sequence, into the genome of the organism, or introducing double-stranded  
CC RNA corresponding to the polynucleotide into the cells of the organism,  
CC thereby inhibiting expression of an antifreeze polypeptide. The  
CC antifreeze protein can be used for the cryopreservation of a cell or  
CC tissue, as a food additive of a frozen food product, in a method for  
CC decreasing the time required to dehydrate a composition, to treat a  
CC disorder characterised by biocrystals associated with disorders such as  
CC gout and kidney stones, to preserve the viability of a molecular biology  
CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
CC to protect a plant from damage due to frost or freezing.  
XX  
SQ Sequence 280 AA;

Query Match 63.8%; Score 884; DB 8; Length 280;  
Best Local Similarity 65.5%; Pred. No. 1.2e-69;  
Matches 186; Conservative 14; Mismatches 52; Indels 32; Gaps 4;





DR N-PSDB; ADM41463.  
XX  
PT New antifreeze proteins and encoding polynucleotides, useful for  
PT modulating cold tolerance in organisms, as food additives, or for  
PT treating tumors or disorders associated with the presence of unwanted  
PT biocrystals (e.g. gout).  
XX  
XX  
PS Claim 7; SEQ ID NO 19; 71pp; English.  
XX  
CC The present sequence is that of an antifreeze protein of tall fescue. The  
CC invention provides forage grass (perennial ryegrass and tall fescue)  
CC antifreeze proteins and the polynucleotides encoding them ADM41458-  
CC ADM41483. The polynucleotides were isolated from tissues taken at  
CC different times of year (winter and spring) and from different parts of  
CC the plants. The polynucleotides can be used to modulate the cold  
CC tolerance of an organism, especially plants, mammals, insects, fungi,  
CC archaea and bacteria. The method involves incorporating an antifreeze  
CC polynucleotide, under the control of a gene promoter sequence, into the  
CC genome of the organism, or introducing double-stranded RNA corresponding  
CC to the polynucleotide into the cells of the organism, thereby inhibiting  
CC expression of an antifreeze polypeptide. The antifreeze protein can be  
CC used for the cryopreservation of a cell or tissue, as a food additive of  
CC a frozen food product, in a method for decreasing the time required to  
CC dehydrate a composition, to treat a disorder characterised by biocrystals  
CC associated with disorders such as gout and kidney stones, to preserve the  
CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
CC patient e.g. tumour tissue, and to protect a plant from damage due to  
CC frost or freezing.  
XX  
SQ Sequence 277 AA;  
Query Match 65.6%; Score 908.5; DB 8; Length 277;  
Best Local Similarity 67.9%; Pred. No. 8e-72;  
Matches 188; Conservative 19; Mismatches 51; Indels 19; Gaps 4;  
QY 5 MAKCCMLVFLGFILOVAGATSWCHDDHLALRGLAENLSGKAVRLRAAMSGASCCSW 64  
DB 1 MAKCCMLLFLVFLPAASAT--SCHPDLRALRGFVGNLNG-GVLLHGAMSGSLCCAW 57  
QY 65 EGVGCTASGRVVALRLP-----KRLGGIIPSSIGELDLHLYLSDNS 109  
DB 58 EGVGCDGTSGRVTALRLPISLKDCKGLKSLNLANLRLVGTIPFISWIGELDLHLYLSDNS 117  
QY 110 LVGEVPSKIQIRLUKSTTDSQSLGMSINMLLHV-SRRTLDEEPTNIGTNNVSGSN 168  
DB 118 LVGKVPNSLQIRLKLATAGSLGMAFANPLHVKGNNRTLDQNTTIGTNNVRSND 177  
QY 169 NVVSGNDNTVSGNNHVSNNNTVGTSDNTVGSNNHVSNGTKHIVTDNNNVSGNDNN 228  
DB 178 NAVSGNDNTVICGNNNTVSGNNNTVSGSDNIVTGSNQVCGTKHITDNNNDVSGNDNN 237  
QY 229 VSGSFHTVSGEHTVSGNNNTVSGSNHIVSGSNKVV 265  
DB 238 VSGSSHVSGSHNTVSGSNNTVSGSNHVSNNKVV 274

RESULT 8  
ADM41483  
ID ADM41483 standard; protein; 281 AA.  
XX  
AC ADM41483;  
XX  
XX 03-JUN-2004 (first entry)  
XX  
DE Perennial ryegrass antifreeze protein.  
XX  
KW Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
KW antigout; litholytic; nephrotropic; cytostatic.  
XX  
OS Lolium perenne.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
Query Match 65.3%; Score 904.5; DB 8; Length 281;  
Best Local Similarity 66.5%; Pred. No. 1.8e-71;  
Matches 187; Conservative 20; Mismatches 51; Indels 23; Gaps 5;  
QY 5 MAKCCMLVFLGFILOVAGATSWCHDDHLALRGLAENLSGKAVRLRAAMSGASCCSW 64  
DB 1 MAKCCMLLFLVFLPAASAT--SCHPDLRALRGFVGNLNGGAL-LRGTWSSGCCDW 57  
QY 65 EGVGCTASGRVVALRLP-----KRLGGIIPSSIGELDLHLYLSDNS 109  
DB 58 EGVGCDGTSGRVTALRLPISLKDCKGLKSLNLANLRLVGTIPFISWIGELDLHLYLSDNS 117  
QY 110 LVGEVPSKIQIRLUKSTTDSQSLGMSINMLLHV-SRRTLDEEPTNIGTNNVSG 164  
DB 118 LVGKVPNSLQIRLKLATAGSLGMAFANPLHVKGNNRTLDQNTTIGTNNVRS 177

FT Region /label= Signal\_peptide  
FT 10..22  
FT /note= "Conserved lipoprotein membrane attachment site"  
FT 53..64  
FT /note= "Conserved Cys-pairs identified N-terminal or  
FT leucine-rich repeats of receptor-like kinases"  
FT 69..125  
FT /note= "Conserved leucine-rich repeats"  
FT 165..279  
FT /note= "7-amino acid sequence repeat region identified in  
FT antifreeze proteins"  
XX WO2004022700-A2.  
XX 18-MAR-2004.  
XX 09-SEP-2003; 2003WO-NZ000199.  
XX 09-SEP-2002; 2002US-0409557P.  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX (WRIG-) WRIGHTSON SEEDS LTD.  
XX Demmer J, Shenk MA, Hall C, Fish SA;  
XX WPI: 2004-248453/23.  
XX N-PSDB; ADM41470.  
XX  
XX New antifreeze proteins and encoding polynucleotides, useful for  
XX modulating cold tolerance in organisms, as food additives, or for  
XX treating tumors or disorders associated with the presence of unwanted  
XX biocrystals (e.g. gout).  
XX  
XX Claim 7; SEQ ID NO 26; 71pp; English.

The present sequence is that of an antifreeze protein of perennial ryegrass. The invention provides forage grass (perennial ryegrass and tall fescue) antifreeze proteins and the polynucleotides encoding them ADM41458-ADM41483. The polynucleotides were isolated from tissues taken at different times of year (winter and spring) and from different parts of the plants. The polynucleotides can be used to modulate the cold tolerance of an organism, especially plants, mammals, insects, fungi, archaea and bacteria. The method involves incorporating an antifreeze polynucleotide, under the control of a gene promoter sequence, into the genome of the organism, or introducing double-stranded RNA corresponding to the polynucleotide into the cells of the organism, thereby inhibiting expression of an antifreeze polypeptide. The antifreeze protein can be used for the cryopreservation of a cell or tissue, as a food additive of a frozen food product, in a method for decreasing the time required to dehydrate a composition, to treat a disorder characterised by biocrystals associated with disorders such as gout and kidney stones, to preserve the viability of a molecular biology reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and to protect a plant from damage due to frost or freezing.





Db 241 EHNVTSGSNNVTSGSNHVSGSNKVVTDG 269  
|||||  
RESULT 3  
ADM41481  
ID ADM41481 standard; protein; 262 AA.  
XX  
AC ADM41481;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Tall fescue antifreeze protein.  
XX  
KW Antifreeze; fescue; cold tolerance; transgenic; plant; antigout;  
KW litholytic; nephrotropic; cytostatic.  
XX  
OS Schedonorus arundinaceus.  
XX  
PH Key Location/Qualifiers  
FT Peptide 1. .23  
FT /label= Signal\_peptide  
FT Region 57. .68  
FT /note= "Conserved Cys-pairs identified N-terminal or  
FT leucine-rich repeats of receptor-like kinases"  
FT Region 73. .114  
FT /note= "Conserved leucine-rich repeats"  
FT Region 147. .261  
FT /note= "7-amino acid sequence repeat region identified in  
FT antifreeze proteins"  
XX  
PN WO2004022700-A2.  
XX  
PD 18-MAR-2004.  
XX  
PF 09-SEP-2003; 2003WO-NZ000199.  
XX  
PR 09-SEP-2002; 2002US-0409557P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (WRIG-) WRIGHTSON SEEDS LTD.  
XX  
PI Demmer J, Shenk MA, Hall C, Fish SA;  
XX  
DR WPI; 2004-248453/23.  
DR N-PSDB; ADM41468.  
XX  
XX New antifreeze proteins and encoding polynucleotides, useful for  
modulating cold tolerance in organisms, as food additives, or for  
treating tumors or disorders associated with the presence of unwanted  
biocrystals (e.g. gout).  
XX  
PS Claim 7; SEQ ID NO 24; 71pp; English.  
XX  
CC The present sequence is that of an antifreeze protein of tall fescue. The  
CC encoding cDNA was isolated from a leaf blade cDNA expression library. The  
CC invention provides forage grass (perennial ryegrass and tall fescue)  
CC antifreeze proteins and the polynucleotides encoding them ADM41458-  
CC ADM41483. The polynucleotides were isolated from tissues taken at  
CC different times of year (winter and spring) and from different parts of  
CC the plants. The polynucleotides can be used to modulate the cold  
CC tolerance of an organism, especially plants, mammals, insects, fungi,  
CC archaea and bacteria. The method involves incorporating an antifreeze  
CC polynucleotide, under the control of a gene promoter sequence, into the  
CC genome of the organism, or introducing double-stranded RNA corresponding  
CC to the polynucleotide into the cells of the organism, thereby inhibiting  
CC expression of an antifreeze polypeptide. The antifreeze protein can be  
CC used for the cryopreservation of a cell or tissue, as a food additive of  
CC a frozen food product, in a method for decreasing the time required to  
CC dehydrate a composition, to treat a disorder characterised by biocrystals  
CC associated with disorders such as gout and kidney stones, to preserve the  
CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
CC patient e.g. tumour tissue, and to protect a plant from damage due to

CC frost or freezing.  
XX  
SQ Sequence 262 AA;  
Query Match 71.7%; Score 992.5; DB 8; Length 262;  
Best Local Similarity 75.0%; Pred. No. 2.8e-79;  
Matches 198; Conservative 22; Mismatches 39; Indels 5; Gaps 3;  
QY 5 MAKCCMLLVPLGFIQVAG-ATWSCHDDHLHALRGLAENLSKGAVRLRAANSASCSS 63  
DB 1 MAKCLMLLSFAFLSAAGTATATPCRRDDLRLRGFAENLGGGALSLRAANSASCSD 60  
QY 64 WEGVGCETASGRVVVALRLPKRGLGGIIPSSIGBLDHLRYLDLGSNLVGEVPKSLQIRLK 123  
DB 61 WEGVGCDGASGRVTALWLPKSGLTGPIPSWICQHLRLYLDLGSNALVGEVPKSLQVQLK 120  
QY 124 SLTTDSQSLGMSGINMLLVH-SSRRTLDEEPNTISGTNNNSVSGSNVSGNDNTVVSGN 182  
DB 121 GLTAAGRS--GFTNMPLHVMNRRLSLDEQNTISGSNNVTSGSNKVVAGNDNTVISGD 177  
QY 183 NNHVSGSNNVTVSGSDNTVSGSNHVSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEHN 242  
DB 178 NNSVSGSNNVTVSGSDNTVSGSNHVSGTNHIVTDNNNVSGNDNNVSGSFHTVSGHNT 237  
QY 243 VSGSNNVTSGSNHVSGSNKVVTD 266  
DB 238 VSGSNNVTSGSNHVSGSNKVVTD 261  
RESULT 4  
ADM41480  
ID ADM41480 standard; protein; 254 AA.  
XX  
AC ADM41480;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Perennial ryegrass antifreeze protein AFP5.  
XX  
KW Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
KW antigout; litholytic; nephrotropic; cytostatic.  
XX  
OS Lolium perenne.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1. .23  
FT /label= Signal\_peptide  
FT Region 57. .68  
FT /note= "Conserved Cys-pairs identified N-terminal or  
FT leucine-rich repeats of receptor-like kinases"  
FT Region 73. .114  
FT /note= "Conserved leucine-rich repeats"  
FT Region 139. .253  
FT /note= "7-amino acid sequence repeat region identified in  
FT antifreeze proteins"  
XX  
PN WO2004022700-A2.  
XX  
PD 18-MAR-2004.  
XX  
PF 09-SEP-2003; 2003WO-NZ000199.  
XX  
PR 09-SEP-2002; 2002US-0409557P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (WRIG-) WRIGHTSON SEEDS LTD.  
XX  
PI Demmer J, Shenk MA, Hall C, Fish SA;  
XX  
DR WPI; 2004-248453/23.  
DR N-PSDB; ADM41467.  
XX  
PT New antifreeze proteins and encoding polynucleotides, useful for









Db 787 AGCTTCATAAAGTATCTGGAAGTCACAATACGGTATCTGGGAGCAACAATACCGGTATCC 846  
Qy 788 GGGAGCAACCATATCGTATCTGGAGCAACAAGTCGTAAACAGATGGTTAATATTCTGTA 847  
Db 847 GGGAGAAACCATGCTGATATCTGGAGCAACAAGTCGTGACAGGAGGTTAATGATATGT 906  
Qy 848 GGTGCA-----GGATTGCTTCCATCTTCCCAAGTTTCAGTGTAGCT 887  
Db 907 AGTGGATTGTTCCATCTCCCTAAGGGATCTCACGTACTTGACAAAGTTCAGTGTAGCA 966  
Qy 888 TACAATCAATAGATGAGACAATCACGTTATGTAATTCAGG-----ATATGGCATACTT 942  
Db 967 CTCAATCACTTGGTGGGCAATCGGTTATGTAACGTCACTGGATATAGCATACGTACTT 1026  
Qy 943 TTCTTTTAATAAAGCTTCCCTTACATAAATAAAAAA 980  
Db 1027 TTCTACTTTAATAAAGACTTTTCCATATAAAAAA 1064

RESULT 12  
US-10-657-852-2  
; Sequence 2, Application US/10657852  
; Publication No. US20040146884A1  
; GENERAL INFORMATION:  
; APPLICANT: Demmer, Jeroen  
; APPLICANT: Shenk, Michael Andrew  
; APPLICANT: Hall, Claire  
; APPLICANT: Fish, Steven A  
; TITLE OF INVENTION: Antifreeze proteins isolated from forage  
; FILE REFERENCE: 11000.1070U  
; CURRENT APPLICATION NUMBER: US/10/657,852  
; PRIOR FILING DATE: 2003-09-09  
; PRIOR APPLICATION NUMBER: 60/409,557  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 959  
; TYPE: DNA  
; ORGANISM: Festuca arundinacea  
US-10-657-852-2

Query Match 27.6%; Score 270.4; DB 22; Length 959;  
Best Local Similarity 74.8%; Pred. No. 1.1e-77;  
Matches 353; Conservative 0; Mismatches 116; Indels 3; Gaps 1;

Qy 383 AGTTTCAGATACGGCTCAAGAGCCTCACACCTGACAGCCAGTCACCTCGGTATGGGTCC 442  
Db 292 AGTTTCAGATATTGCTCAAGGGCTCACACCGCTGGCGGTTCACTGGGTAAAGGGTTC 351  
Qy 443 ATTAACATGCTATTGATGTGAGCAGTAG---AAGAACGCTCGATGAAGAACCAATACA 499  
Db 352 ACTAACATGCCATTACATGTGAAGTCTAGCCAAAGGAACACTCGACGAAGAACACAATACA 411  
Qy 500 ATATCAGGAGCAACAATAGTTGATCAGGAGCAACAATGTTGTTTCCGGGAATGAC 559  
Db 412 ATAACCTGGGATCAACAATACTGTCAAAATCCGGGAGCAACAATGTTGTTTCTGGGAACGAT 471  
Qy 560 AACACGGTCTGATCTGCGGAATACACCATGTGCTCGGAGCAACAACACTGTTGTAACCT 619  
Db 472 AACACTGTATATCCGGGAAACAACACGTCGTGTCGGGAGGCCAACACCGTCGTATTT 531  
Qy 620 GGAAGTGACAATACTGTAGTCTGATGCAACCATGTGCTATCAGGGAACAAGCATATTGTT 679  
Db 532 GGGGTTGACAAATTCCTAAGCGGTAGCAACCATGTGCTATCTGGGAACCACTGTGCGT 591  
Qy 680 ACTGATAACAATAATGTTGATCCGGGAACGACAATAATGTGTCTGGAAGCTTCCATCT 739  
Db 592 ACTGACAACAAGAATGCCGTATCCGGGGACCAATACTGTATCTGGAAGCCAAAATACC 651  
Qy 740 GTATCAGGGGAGCACAAATACCGTATCCGGGAGCAACAATACTGTATCCGGGAGCAACCAT 799

Db 652 GTATCCGGGAACCAACCATATCATATCTCGAGCCACAGTACCATATATCCGGGAACCAAT 711  
Qy 800 ATCGTATCTGGAGCAACAAGTCGTAAACAGATGGTTAATATTCTCTAGGTG 851  
Db 712 ACCGTATCCGGGAGCAACAATTTCTGTATCTGGGAACAACAATATTGTATCTG 763

RESULT 13  
US-10-657-852-1  
; Sequence 1, Application US/10657852  
; Publication No. US20040146884A1  
; GENERAL INFORMATION:  
; APPLICANT: Demmer, Jeroen  
; APPLICANT: Shenk, Michael Andrew  
; APPLICANT: Hall, Claire  
; APPLICANT: Fish, Steven A  
; TITLE OF INVENTION: Antifreeze proteins isolated from forage  
; FILE REFERENCE: 11000.1070U  
; CURRENT APPLICATION NUMBER: US/10/657,852  
; PRIOR FILING DATE: 2003-09-09  
; PRIOR APPLICATION NUMBER: 60/409,557  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 841  
; TYPE: DNA  
; ORGANISM: Festuca arundinacea  
US-10-657-852-1

Query Match 26.8%; Score 262.2; DB 22; Length 841;  
Best Local Similarity 73.5%; Pred. No. 5.3e-75;  
Matches 349; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

Qy 380 AAAAGTTTCAGATACGGCTCAAGAGCCTCACACCTGACAGCCAGTCACCTCGGTATGGGT 439  
Db 262 AAGAGCTTCAGATATTGCTCAAGGGCTCACCGCGCTGCGGTCCTGTTCTGTTAGGCG 321  
Qy 440 TCCATTAAACATGCTATTGATGTGAGCAGTAG---AAGAACGCTCGATGAAGAACCAAT 496  
Db 322 TTCCTCACATGCCATTACATGTGAAGCCTAGCAAGGAACACTCGACGAAGACCACAAT 381  
Qy 497 ACAATATCAGGAGCAACAATAGTTGATCAGGAGCAACAATGTTCTTCCGGGAAT 556  
Db 382 ACAATAAAGTGGATCAACAATACTGTTCAGATCCGGGAGCAACAATGTTGTTCTGGGAAC 441  
Qy 557 GACAACACCGTCTGATCTGGGAATAACAACCATGTGTCTGGGAGCAACAACACTGTTGTA 616  
Db 442 GATAACATGTCATATCCGGGAACAACAACGTCGTGTCGGGAGCCACAACCCGTCGTA 501  
Qy 617 ACTGGAAGTGACAATACTGTAGTTGGTAGCAACCATGTGCTATCAGGGAACAAGCATATT 676  
Db 502 TTTGGGGTGACAACCTTCATAAGTGAAGTTACCATGTGCTATCTGGGAACCACTATGTT 561  
Qy 677 GTTACTGATAACAATAATGTTGATCCGGGAACGACAATAATGTGTCTGGAAGCTTCCAT 736  
Db 562 GTGACTGACAACAAGAATGCCGTATCCGGGGAGCCAACTACTGTATCTGGAAGCCAAAAT 621  
Qy 737 ACTGTATCAGGGAGCACAAATACCGTATCCGGGAGCAACAATACTGTATCTCCGGAGCAAC 796  
Db 622 ACCGTATCCGGGAACCAACAGATCGTATCTGGGAGCCACAGTACCGGTATCCGGGAACCA 681  
Qy 797 CATATCGTATCTGGGAGCAACAAGTCCGTAACAGATGGTTAATATTCTGTAGGTG 851  
Db 682 ATACGGTATCTGGGAGAAACAATTCGGTATATGGGAACAACAATATTGTATCTG 736

RESULT 14  
US-10-963-77805  
; Sequence 77805, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:





```

Db      837  CTGGAGCAACAATACCGTATCGGGAGAAACCATGTCGTATCTGGAGCAACAAATCG 896
QY      825  TAACAGATGTTAATATTCTGTAGGTGCA-----GGATGCTTCC 864
Db      897  TGACAGAGGTTAATGATATGTTAGTGGATTGTTCCATCTCCCTTAAGGGATCTCACGT 956
QY      865  ATCTTCCCAAGTTCAGTGTAGCTTACAAATCAATAGATGAGAGACAATCAGCTTATGTAAC 924
Db      957  ACTTGACAGAGTTGAGTTGAGCACTCAATCATTGTTGGGACAATCGGTTATGTAAC 1016
QY      925  TCA--GGATATGCCATATTTTCC 947
Db      1017 TCATGGATATATCATACCTTTCCT 1041

RESULT 9
US-10-657-852-13
; Sequence 13, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Lolium perenne
US-10-657-852-13

```

```

Query Match      42.8%; Score 419; DB 22; Length 1083;
Best Local Similarity 70.4%; Pred. No. 1.2e-126;
Matches 702; Conservative 0; Mismatches 205; Indels 90; Gaps 7;

QY      33  CCATGCTGAATACATACGCAAGATGTGATGCTGTGCTTCTTGGGGTTCACTTGC 92
Db      45  CCAGAACTTAATCCATGCGGAAATGTTGGTGTGCTGTCTTCTTGGTGTCTTCTTGC 104
QY      93  AGTGGCAGGCAACGTCGTGTGTCGCCACACGACACCTCCACGCGTTGAGGGCC 152
Db      105  TGGCCATGAGCGGAC-----GTCTGCCACTTGGATGACCTCGCGCGCTCGGGGCT 158
QY      153  TCGCTGAGAACCTTAAGCGGCAAGAGCGCTCGCTCCGCGCGCATGGTCCGGCGCT 212
Db      159  TTGTGGGAACCTCAATGCG---GGGGTGCCCTTCTCGTGGACATGGTCTGGCTCT 215
QY      213  CATGCTGACGTGGGAAGTGTGGATGCGAAACAGACGCGCGCGCTGTGGCGTTGC 272
Db      216  CATGCTGCGATTGGGAAGGTGTGGGTGCGATGGTACAAAGCGCGCGCTCACGGGCT 275
QY      273  GGCTCC-----CCAAGCGG 287
Db      276  GGCTTCCGATTAAGCTCGAGGACTGCGGTAAGCTCAAGTCGCTCAACCTTGCACGAAA 335
QY      288  GCTTGGAGGGATCATCCCATCGTTCGATTGGTGAGCTTGATCACTTTCGCTATTGGATC 347
Db      336  GATTGGTGGACCATCCCGTGTGGATTGGTGGAGCTTGACCAACCATGCTACTTGGTTC 395
QY      348  TCTCGGGTAATTAATTATTTGGTTGGGAGGTACC-----AAAAAGTTTGCAGATAC 395
Db      396  TCTCGGATAATTCAATTGGTTGGTAAGGCACCAACCAATAGTTTGCACAATAGTTTGCAGATAA 455
QY      396  GGCCTCAAGAGCCTCACCATGACAGCCAGTCACTCGGTATGGGTCCATTAAATGCTAT 455

```

```

Db      456  GACTCAAGGCGCTCGCCACCGCTGTCGTTCACTAGGTATGGCTTTCGCTAAATGCCAT 515
QY      456  TGCATGTGA---GCAGTAGAAGAACGCTCGATGAAGAACCAATATACATATCAGGACCA 512
Db      516  TGCATGTGAAGGGAACCGAAGAACCCCTCGAGAACAAACAAATACATAATACATGGGACCA 575
QY      513  ACATATAGTCTTGGATCAGGAGCAACAATGTTGTTTCCGGGAATGACCAACACGTCGTAT 572
Db      576  ACAACACTGTGTAGATCTGGGAACGACAATGCTGTTTCTGGGAACGACCAACACTGTCTAT 635
QY      573  CTGGGAATAACCAACCATGTCCTGGGAGCAACAACTGTTGTAACCTGGAAGTGACAATA 632
Db      636  GTGGGNAACAACAACACTGTCCTGGGAGCAACAACACCATATCAATACTGGCAGTGACATA 695
QY      633  CTGTAGTTGGTAGCAACCATGTCGTATCAGGGACAAAGCATATTTGTACTGTATAACAATA 692
Db      696  TCGTAACTGGCAGCAACCATATTTGTATGTGGGACCAACATATCAATACTGATAACAACA 755
QY      693  ATGTTGTATCCGGGAACGACAATAATGTCCTGGAGAGCTTCCATCTATCTAGGGGAGC 752
Db      756  ATGACGTATCCGGCAATGATATGATGTCCTGGGAGCTTCCATCTATCTCGGGAGCC 815
QY      753  ACAATACCGTATCCGGGAGCAACAATACTGTATCCGGGAGCAACCATATCTGTATCTGGGA 812
Db      816  ACAATACTGTATCTGGAGTAACAACACTGTATCTGGAAGCAACCATGTCGTATCTGGAA 875
QY      813  GCAACAAAGTCGTAAACAGATGTTAATATTTCTGTAG-----TGC 852
Db      876  GCAACAAAGTCGTGACAGGAGATGAATGATTTGTACGGGATTTGCTTCCATCTTCTCTAA 935
QY      853  AGGATGCTTCCATCTTCCAAAGTTCAAGTTCAGTGTAGCTTACATCAATAGATGGAGACAATCA 912
Db      936  AGGAGCTCTCACCTTAGTCCAAAGTTGCGTSCAGCTCACAATCACTTGGTAGGACAATCG 995
QY      913  CGTTATGTAACCTTCA--GGATATGGCATACTTTTCTT 948
Db      996  AGTTATGTAACCTTCAATGATATAGCATCATCTTCCCT 1032

```

```

RESULT 10
US-10-657-852-5
; Sequence 5, Application US/10657852
; Publication No. US20040146884A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; TITLE OF INVENTION: grasses and methods for their use.
; FILE REFERENCE: 11000.10700
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Festuca arundinacea
US-10-657-852-5

```

```

Query Match      42.6%; Score 417.4; DB 22; Length 1084;
Best Local Similarity 70.3%; Pred. No. 4e-126;
Matches 701; Conservative 0; Mismatches 206; Indels 90; Gaps 7;

QY      33  CCATGCCCTGAATACATACGCAAGTGTGCAATGCTGCTGTCTTCTTGGGGTTCACTTGC 92
Db      46  CCAGAACTTAATCCATGCGAATGTTGGTGTGCTGTCTTCTTGGTGTCTTCTTCTTGC 105
QY      93  AGGTGGCAGGAGCAACGTCGTGTGTCGCCACCAACGACCTCCACGCGTTGAGGGGCC 152

```





QY	72	TCTTCTTTGGGGTTTCATCTTGAGGTGGCAGGA---GCAACGTCGTGTGTCGTGCGCACCCACG	128
DB	104	TCTCTCTTCGGGTTCTCTTGTTCGGTGGCCGCA CGGCGACGGCGACGCATGCCACCGCG	163
QY	129	ACGACTCTCCACGGGTGAGGGGCTTCGTGAGAACCTAAGCGGCAAAAGAGCGCTCCGCG	188
DB	164	ATGACCTTCGGCGCTGCGGGGCTTCGTGTGAGAACCTGGGCGGCGCGCAATCAGCC	223
QY	189	TCCGGCGCGATGTCGCGGCGCTCATGCTCAGCTGGGAGGTCTGTGGATCGGAAACAG	248
DB	224	TCCGGCGCGGTGTCGAGCGCTCATGCTGCGATTGGGAAGGCGTTGCTGCGACGGTG	283
QY	249	CAAGCGGCGCGTCTGTCGCGTTGCGGCTCCCCAAGCGCGGCTTCGGAGGGATCATCCCAT	308
DB	284	CCAGGGCGCGTGTACGGCTTTGTGGCTCCCGAGGAGCGGCTCACGGGGCCAATCCCGT	343
QY	309	CGTCGATTGGTGAGCTTGATCAACCTTCGCTATTGGATCTCTCGGGTAACTCAATGGTTG	368
DB	344	CATGGAATTTGTGAGCTTCACCACTTACGCTACTTCGATCTTTACGATTAATGCAATGGTTG	403
QY	369	GGGAGGTACCAAAAGTTTGCAGATACGGCTCAAGAGCCTCACCACTGACAGCCAGTCAAC	428
DB	404	GCGAGGTACCCAAGAATCTGACGATCACAGCTCAAAAGGCATCAC	447
QY	429	TCGGTATGGGTTCCATTAACTATGCTATTGTCATGTCA---GCAGTAGAGAAAGCGCTCGATG	485
DB	448	-----AACATGCCATTGCAATGTCATGTCGATTCGCTAAACAGAGATCACTTCGACG	490
QY	486	AAGAACCAATAACAATATCAGGAGCAACAACATAGTTGTTGGATCAGGGAGCAACAATGTTG	545
DB	491	AGCAGCCCAATACAATTTCTGGGAGCAACAATACTGTTCAGATCCGGGAGCAAAAATGTTG	550
QY	546	TTTTCGGGAATGACAACAACGGTTCGTAATCTGGGAATAACAACATGTCTCGGGAGCAACA	605
DB	551	TTGCTCGGAATGACAACAACCGTTCATCTGGGGACAACAATAAGTTGTCTGGGAGCAACA	610
QY	606	ACACTGTTTAACCTGGAGTGACATACTGTAGTTGTTAGCAACCATGTCGTATCAGGGA	665
DB	611	ACACTGTCGTAAGTGGGAATGACAATACCGTAACCGGAGCAACCATGTCGTATCAGGGA	670
QY	666	CAAAAGCATATGTTACTGATAACAATAATGTTGTATCCGGGAACGACAATAATGTGCTG	725
DB	671	CAAAACCATATCGTTACAGACAACAACATAACGATTCGGGAACGATAATATGATTCG	730
QY	726	GAAAGTTCCTCATCTGTATCAGGGGAGCAATAACCGTATCCGGGAGCAACAATCTGTAT	785
DB	731	GGAGCTTTTCATACCGTATCCGGGGGGCAACAATACTGTCTCCGGGAGCAACAATACCGTAT	790
QY	786	CCGGGAGCAACCATATCGTATCTCGGGAGCAACAAGTCGTACAGATGGTTAATATCTG	845
DB	791	CTGGGAGCAACCAACCGTTGTAATCTGGAAGCAACAAGTCGTGACAGACGCTTAATGATCTG	850
QY	846	TAGTGCAGGATTGCTTCCATCTT-----CCCAAGTTCAGTGTAGC	886
DB	851	TCAGCGCATGATGTTTCCACCTTAACCTGAGCTCAGTTCCTGTCGAAGTTTCACTGTACC	910
QY	887	TTACAATCAATAGATGGAGACAATCACGCTTATGTAACTTCA--GGATATGGCATACTTTTC	945
DB	911	TCACAGTCAGTTGGTTCGCTTCAATCCGGTTATGTAACCTTCATGGATATACCATACTTTTC	970
QY	946	---CTTTAAATAAAGCTTCCCTTTTACATAAAAAAAA	980
DB	971	CTACTATATATAAAATTTCCCTTTACATAAAAAAAA	1008

## RESULT 5

US-10-657-852-12  
; Sequence 12, Application US/10657852  
; Publication No. US20040146894A1  
; GENERAL INFORMATION:  
; APPLICANT: Demmer, Jeroen  
; APPLICANT: Shank, Michael Andrew





```

Db      1  GCTTGCATTCGAATCAAGGTTCTTGTTCATCCATGCTGAATACATGGCAAGTGTG 60
Qy      61  CATGCTGCTGCTTCTTCTTGGGTTTCATCTTGAGGTGGCAGAGCAACGTCGTGTCGTG 120
Db      61  CATGCTGCTGCTTCTTCTTGGGTTTCATCTTGAGGTGGCAGAGCAACGTCGTGTCGTG 120
Qy     121  CCACACGACGACCTCCAGCGCTTGAGGGGCTCGCTGAGACCTTAAGCGGCAAGAGGC 180
Db     121  CCACACGACGACCTCCAGCGCTTGAGGGGCTCGCTGAGACCTTAAGCGGCAAGAGGC 180
Qy     181  CGTCCGCTCCGCGCGCATGTCGCGCGCTCATGCTGCAGCTGGAAGGTGTGGATG 240
Db     181  CGTCCGCTCCGCGCGCATGTCGCGCGCTCATGCTGCAGCTGGAAGGTGTGGATG 240
Qy     241  CGAAACAGCAAGCGCGCGCTGTCGGCGTTGCGGCTCCCAAGCGGCGCTTGGAGGAT 300
Db     241  CGAAACAGCAAGCGCGCGCTGTCGGCGTTGCGGCTCCCAAGCGGCGCTTGGAGGAT 300
Qy     301  CATCCATCGTCGATGCTGAGCTTCATCACCCTCGCTATTTGGATCTCTCGGTAATTC 360
Db     301  CATCCATCGTCGATGCTGAGCTTCATCACCCTCGCTATTTGGATCTCTCGGTAATTC 360
Qy     361  ATTGGTTGGGAGGTACCAAAAGTTTGAGATACCGCTCAAGAGCTCACCACTGACAG 420
Db     361  ATTGGTTGGGAGGTACCAAAAGTTTGAGATACCGCTCAAGAGCTCACCACTGACAG 420
Qy     421  CCAGTCACCTCGGTATGGGTTCCATTAACTATGCTATTTGATGAGAGCTTGAAGACGCT 480
Db     421  CCAGTCACCTCGGTATGGGTTCCATTAACTATGCTATTTGATGAGAGCTTGAAGACGCT 480
Qy     481  CGATGAGAACCAATATACATATCAGGACCAACATAGTTGATCAGGAGCAACAA 540
Db     481  CGATGAGAACCAATATACATATCAGGACCAACATAGTTGATCAGGAGCAACAA 540
Qy     541  TGTGTTTCCGGAATGACAAACAGCTCGTATCTGGGAATAACCACTGTCTGGGAG 600
Db     541  TGTGTTTCCGGAATGACAAACAGCTCGTATCTGGGAATAACCACTGTCTGGGAG 600
Qy     601  CAACAACTGTGTTAACTGGAAGTGACAATACTGTAGTTGGTAGCAACCATGTCTATC 660
Db     601  CAACAACTGTGTTAACTGGAAGTGACAATACTGTAGTTGGTAGCAACCATGTCTATC 660
Qy     661  AGGGAACAAGCATTTGTTACTGATTAACAATAATGTTGATCCGGGAACGACAATAAGT 720
Db     661  AGGGAACAAGCATTTGTTACTGATTAACAATAATGTTGATCCGGGAACGACAATAAGT 720
Qy     721  GTCTGGAAGCTTCCATCTGATCAGGAGGACAAATACCGTATCCGGGAGCAACAATAC 780
Db     721  GTCTGGAAGCTTCCATCTGATCAGGAGGACAAATACCGTATCCGGGAGCAACAATAC 780
Qy     781  TGTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTAAACAGATGGTTAATA 840
Db     781  TGTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTAAACAGATGGTTAATA 840
Qy     841  TTCTGTAGTGCAGGATTCGTTCCATCTTCCCAAGTTTCAGTGTAGCTTCAATCAATAGA 900
Db     841  TTCTGTAGTGCAGGATTCGTTCCATCTTCCCAAGTTTCAGTGTAGCTTCAATCAATAGA 900
Qy     901  TGGAGCAATACAGTTATGTAATCTCAGGATATGGATCTTTTCCCTTTAAATAAGCTT 960
Db     901  TGGAGCAATACAGTTATGTAATCTCAGGATATGGATCTTTTCCCTTTAAATAAGCTT 960
Qy     961  CCCTTTTACATAAAAAAAA 980
Db     961  CCCTTTTACATAAAAAAAA 980

```

RESULT 2  
 US-10-657-852-4  
 ; Sequence 4, Application US/10657852  
 ; Publication No. US20040146884A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Demmer, Jeroen

```

; APPLICANT: Sherk, Michael Andrew
; APPLICANT: Hall, Claire
; APPLICANT: Fish, Steven A
; TITLE OF INVENTION: Antifreeze proteins isolated from forage
; FILE REFERENCE: 11000.1070U
; CURRENT APPLICATION NUMBER: US/10/657,852
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/409,557
; PRIOR FILING DATE: 2002-09-09
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Festuca arundinacea
; US-10-657-852-4

Query Match      92.3%; Score 904.4; DB 22; Length 996;
Best Local Similarity 96.2%; Pred. No. 8.6e-287;
Matches 949; Conservative 0; Mismatches 31; Indels 6; Gaps 2;

Qy      1  GCTTGCATTCGAATCAAGGTTTCTTGTTCATTCATGCTGAATACATGGCAAGTGTG 60
Db      7  GCTTGCATTCGAATCAAGGTTTCTTGTTCATTCATGCTGAATACATGGCAAGTGTG 66
Qy     61  CATGCTGCTGCTTCTTCTTGGGTTTCA---TCTTGCAGTGGCAGGAGCAACGTCGTGTC 117
Db     67  CATGCTGCTGCTTCTTCTTGGGCTTTCATCTCTTGCAGTGGCGGAGCAACGTCGTGTC 126
Qy    118  GTGCCACCAACGACGACCTCCACGCTTCAGGGGCTCGCTGAGAACCTTAAGCGCAAGG 177
Db    127  GTGCCACCAACGACGACCTCCGCGCATTCAGGGGCTTCGCGGAGAACCTTAAGCGCAAGG 186
Qy    178  AGCCGTCCGCTCCGCGCGCATGTCGCGGCTCATGCTGCAGCTGGGAAGTGTGG 237
Db    187  AGCCGTCCGCTCCGCGCGCATGTCGCGGCTCATGCTGCAGCTGGGAAGTGTGG 246
Qy    238  ATGCGAAACAGCAAGCGCGCTCGTGGCGTTCGCGCTCCCAAGCGGCGCTTGGAGG 297
Db    247  ATGCGAAACAGCAAGCGCGCTCGCGCTTCGCGCTCCCAAGCGGCGCTTGGAGG 306
Qy    298  GATCATCCATCGTCGATTTGGTGAAGTTCATCACTTCGCTATTTGGATCTCTCGGGTAA 357
Db    307  GATCATCCATCGTCGATTTGGTGAAGTTCATCACTTCGCTATTTGGATCTCTCGGGTAA 366
Qy    358  TTCAATTGTTGGGAGGTACCAAAAGTTTGCAGTACGGCTCAAGAGCTCAACACTGA 417
Db    367  TTCAATTGTTGGGAGGTACCAAAAGTTTGCAGTACGGCTCTAGAGCTCTCCACTGA 426
Qy    418  GAGCAGTCACTCGGTATGGGTTCCATTAAATGCTATTTGATGAGCAGT---AGAG 474
Db    427  TGCGCAGTCACTCGGTATGGGTTCCATTAAATGCTATTTGATGAGCAGTAAACAGAG 486
Qy    475  AACGCTTCGATGAAGAACCAATAATACAGGACCAACATAGTTTGGATCAGGGAG 534
Db    487  AACCTTCGATGAAGAACCAATAATACAGGACCAACATAGTTTGGATCAGGGAG 546
Qy    535  CAACAATGTTGTTCCGGGAATGACAAACGTCGTATCTGGGAATAACCAACTGTGTC 594
Db    547  CAACAATGTTGTTCCGGGAATGACAAACGTCGTATCTGGGAATAACCAACTGTGTC 606
Qy    595  TGGGAGCAACACACTGTTGTAACCTGGAAGTGAACATCTAGTTGGTAGCAACCATCT 654
Db    607  TGGGAGCAACACACTGTTGTAACCTGGAAGTGAACATCTAGTTGGTAGCAACCATCT 666
Qy    655  CGTATCAGGACAAAGAGCATATTTGTTACTGATAAATATGTTGATCCGGGAACACAA 714
Db    667  CGTATCAGGACAAAGAGCATATTTGTTACTGATAAATATGTTGATCCGGGAACACAA 726
Qy    715  TAATGTGCTCGGAAGCTTCCATCTGATCAGGGAGCAACATACCGTATCCGGGAGCAA 774
Db    727  TAATGTGCTCGGAGCTTCCATCTGATCAGGGAGCAACATACCGTATCTGGGAGCAA 786

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2005, 19:05:46 ; Search time 760 Seconds  
(without alignments)

8980.980 Million cell updates/sec

Title: US-10-657-852B-3

Perfect score: 980

Sequence: 1 gcttgattcccaatcaaggt.....ccctttacataaaaaaaaaa 980

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 8443130 seqs, 3482420727 residues

Total number of hits satisfying chosen parameters: 16886260

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	980	100.0	980	22	US-10-657-852-3
2	904.4	92.3	996	22	US-10-657-852-4
3	541.6	55.3	1006	22	US-10-657-852-11
4	503.6	51.4	1008	22	US-10-657-852-10
5	492.8	50.3	1007	22	US-10-657-852-12
6	451.2	46.0	1230	22	US-10-657-852-6
7	443.6	45.3	1212	22	US-10-657-852-7
					Sequence 3, Appli
					Sequence 4, Appli
					Sequence 11, Appl
					Sequence 10, Appl
					Sequence 12, Appl
					Sequence 6, Appli
					Sequence 7, Appli

8	421.8	43.0	1246	22	US-10-657-852-9	Sequence 9, Appli
9	419	42.8	1083	22	US-10-657-852-13	Sequence 13, Appl
10	417.4	42.6	1084	22	US-10-657-852-5	Sequence 5, Appli
11	394.4	40.2	1064	22	US-10-657-852-8	Sequence 8, Appli
12	270.4	27.6	959	22	US-10-657-852-2	Sequence 2, Appli
13	262.2	26.8	841	22	US-10-657-852-1	Sequence 1, Appli
14	79.4	8.1	3285	19	US-10-437-963-77805	Sequence 77805, A
15	79.4	8.1	4536	17	US-10-260-238-1496	Sequence 1496, Ap
16	67.2	6.9	495335	22	US-10-737-082-12	Sequence 12, Appl
17	67.2	6.9	495635	22	US-10-765-790-10	Sequence 10, Appl
18	67.2	6.9	705636	22	US-10-737-082-30	Sequence 30, Appl
19	67.2	6.9	705636	22	US-10-765-790-30	Sequence 30, Appl
20	50.8	5.2	2883	19	US-10-437-963-4582	Sequence 4582, Ap
21	50.6	5.2	1912	17	US-10-437-963-47843	Sequence 568, App
22	50.6	5.2	2037	19	US-10-437-963-47843	Sequence 47843, A
23	45.4	4.6	2498	13	US-10-027-632-102064	Sequence 102064, A
24	45.4	4.6	2498	13	US-10-027-632-102065	Sequence 102065, A
25	45.4	4.6	2498	13	US-10-027-632-102066	Sequence 102066, A
26	45.4	4.6	2498	17	US-10-027-632-102064	Sequence 102064, A
27	45.4	4.6	2498	17	US-10-027-632-102065	Sequence 102065, A
28	45.4	4.6	2498	17	US-10-027-632-102066	Sequence 102066, A
29	44.8	4.4	1614	19	US-10-437-963-79739	Sequence 79739, A
30	44.8	4.4	3453	14	US-10-101-464A-861	Sequence 861, App
31	44.8	4.4	3453	21	US-10-864-252-861	Sequence 861, App
32	44.6	4.6	2589	19	US-10-437-963-47358	Sequence 47358, A
33	43.8	4.5	235070	13	US-10-087-192-1590	Sequence 1990, Ap
34	43.4	4.4	485	20	US-10-425-115-155868	Sequence 155868, A
35	43.4	4.4	1504	20	US-10-719-993-27	Sequence 27, Appl
36	43.4	4.4	1569	20	US-10-719-993-26	Sequence 26, Appl
37	43.4	4.4	37305	20	US-10-719-993-6767	Sequence 6767, Ap
38	42.8	4.4	2379	19	US-10-437-963-7515	Sequence 7515, Ap
39	42.4	4.3	632	18	US-10-424-599-138064	Sequence 138064, A
40	42.4	4.3	990	20	US-10-425-115-105088	Sequence 105088, A
41	42.4	4.3	2485	18	US-10-425-114-24982	Sequence 24982, A
42	42.2	4.3	2268	19	US-10-437-963-39626	Sequence 39626, A
43	42.2	4.3	2423	20	US-10-425-115-3469	Sequence 3469, Ap
44	41.8	4.3	492	19	US-10-767-701-22961	Sequence 22961, A
45	41.8	4.3	3276	19	US-10-437-963-30048	Sequence 30048, A

ALIGNMENTS

RESULT 1  
US-10-657-852-3  
; Sequence 3, Application US/10657852  
; Publication No. US20040146884A1  
; GENERAL INFORMATION:  
; APPLICANT: Demmer, Jercoen  
; APPLICANT: Shenk, Michael Andrew  
; APPLICANT: Hall, Claire  
; APPLICANT: Fish, Steven A  
; TITLE OF INVENTION: Antifreeze proteins isolated from forage  
; FILE REFERENCE: 11000.1070U  
; CURRENT APPLICATION NUMBER: US/10/657,852  
; CURRENT FILING DATE: 2003-09-09  
; PRIOR APPLICATION NUMBER: 60/409,557  
; PRIOR FILING DATE: 2002-09-09  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 980  
; TYPE: DNA  
; ORGANISM: Festuca arundinacea  
US-10-657-852-3

Query Match 100.0%; Score 980; DB 22; Length 980;  
Best Local Similarity 100.0%; Pred. No. 9,9e-312;  
Matches 980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTGATTCCCAATCAAGGTTCCTTTCCTCAATCCATGCTGAATACATGCAAGTGTTC 60  
|||||



**This Page Blank (uspto)**

**~~This Page Blank (uspto)~~**



```
Db 297 CCTGAGTGTGCTGTTGAATGTTAGCTCTCCCGCTATATTGGGAACCTCTCTCTTTTGGC 356
Qy 337 CTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTTCTTGAATGCTTTGAGTGGGAGAAATTCAGCAGAGTT 407

RESULT 13
US-10-101-464A-39
; Sequence 39, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIORITY FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-39

Query Match 3.9%; Score 38.2; DB 4; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 217 CTGAGCTGGGAAGTGTGGGATCGAAACAGCAAGCGCGCTCGTGGCGTTCGGCT 276
Db 237 CTGCTGTGGACGGAATTCGTGAGGCCAAACAGCAAGAGTGTTCATTTCTCT 296
Qy 277 CCCAAGCGCGCTTGGAGGATCATCCCATCGTTCGATTCGATTCACCTTCG 336
Db 297 CCTGAGTGTGTTGAATGCTCTCCCGCTATATTGGGAACCTCTCTCTTTTGGC 356
Qy 337 CTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTTCTTGAATGCTTTGAGTGGGAGAAATTCAGCAGAGTT 407

RESULT 14
US-10-101-464A-421
; Sequence 421, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIORITY FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
US-10-101-464A-421

Query Match 3.9%; Score 38.2; DB 4; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 217 CTGAGCTGGGAAGTGTGGGATCGAAACAGCAAGCGCGCTCGTGGCGTTCGGCT 276
Db 237 CTGCTGTGGACGGAATTCGTGAGGCCAAACAGCAAGAGTGTTCATTTCTCT 296
Qy 277 CCCAAGCGCGCTTGGAGGATCATCCCATCGTTCGATTCGATTCACCTTCG 336
Db 297 CCTGAGTGTGTTGAATGCTCTCCCGCTATATTGGGAACCTCTCTCTTTTGGC 356
Qy 337 CTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTTCTTGAATGCTTTGAGTGGGAGAAATTCAGCAGAGTT 407

RESULT 15
US-09-621-976-15639
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match 3.9%; Score 38.2; DB 4; Length 505;
Best Local Similarity 12.3%; Pred. No. 0.082;
Matches 54; Conservative 188; Mismatches 196; Indels 1; Gaps 1;

Qy 184 CCGCTCGCGCCGATGGTCCGGCGCTCATCTGCAGCTGGGAAGTGTGGGATGCGA 243
Db 6 SSRYTSSKRYGKYSSMYSSRSMKYAWGRKYGTSGRGGSRGMMCKWGYRYSY 65
Qy 244 AACAGCAAGCGCGCTCGTGGCGTTCGGCTCCCAAGCGCGCTTGGAGGATCAT 303
Db 66 WGYKWSKSKMKYSGHGTSSKSKRYRYSKRTTCKYRWSKSKWMMKRRKMY 124
Qy 304 CCCATCGTGTGATTCGATTCACCTTCGATTCGATTCGATTCGATTCATTT 363
Db 125 YRMKCYSCASYSYRRCRYTGTMTGMYCKRMCKSKSTRYMTRYTRYWMTGACYGS 184
Qy 364 GTTGGGAGTACCAAAAAGTTTCAGATACGGCTCAAGAGCTCACCCTGACAGCCA 423
Db 185 KMSCKGRSKYSGYKRYGKYTCTSKYKSKMYKSKMYKSKMYKSKMYKSKMYKSKMY 244
Qy 424 GTCATCGTGTGATTCATTAACATTCATTCATTCATTCATTCATTCATTCATTC 483
Db 245 KKSYYTCKRSYYTYRTSTSKGWTGKSKRWSYTWGSKSKRWSYTWGSKSKRWSYTW 304
Qy 484 TGAAGAACCAAAATACAATATACAGGACCAACAATAGTGTTCGATTCAGGAGCAACA 543
Db 305 MSWCMCAWMAWMSRMAWMSRMAWMSRMAWMSRMAWMSRMAWMSRMAWMSRMAWMSR 364
```

```
Db 297 CCTGAGTGTGCTGTTGAATGTTAGCTCTCCCGCTATATTGGGAACCTCTCTCTTTTGGC 356
Qy 337 CTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTTCTTGAATGCTTTGAGTGGGAGAAATTCAGCAGAGTT 407

RESULT 13
US-10-101-464A-39
; Sequence 39, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIORITY FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-39

Query Match 3.9%; Score 38.2; DB 4; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 217 CTGAGCTGGGAAGTGTGGGATCGAAACAGCAAGCGCGCTCGTGGCGTTCGGCT 276
Db 237 CTGCTGTGGACGGAATTCGTGAGGCCAAACAGCAAGAGTGTTCATTTCTCT 296
Qy 277 CCCAAGCGCGCTTGGAGGATCATCCCATCGTTCGATTCGATTCACCTTCG 336
Db 297 CCTGAGTGTGTTGAATGCTCTCCCGCTATATTGGGAACCTCTCTCTTTTGGC 356
Qy 337 CTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTTCTTGAATGCTTTGAGTGGGAGAAATTCAGCAGAGTT 407

RESULT 14
US-10-101-464A-421
; Sequence 421, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIORITY FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
US-10-101-464A-421

Query Match 3.9%; Score 38.2; DB 4; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 217 CTGAGCTGGGAAGTGTGGGATCGAAACAGCAAGCGCGCTCGTGGCGTTCGGCT 276
Db 237 CTGCTGTGGACGGAATTCGTGAGGCCAAACAGCAAGAGTGTTCATTTCTCT 296
Qy 277 CCCAAGCGCGCTTGGAGGATCATCCCATCGTTCGATTCGATTCACCTTCG 336
Db 297 CCTGAGTGTGTTGAATGCTCTCCCGCTATATTGGGAACCTCTCTCTTTTGGC 356
Qy 337 CTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTTCTTGAATGCTTTGAGTGGGAGAAATTCAGCAGAGTT 407

RESULT 15
US-09-621-976-15639
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match 3.9%; Score 38.2; DB 4; Length 505;
Best Local Similarity 12.3%; Pred. No. 0.082;
Matches 54; Conservative 188; Mismatches 196; Indels 1; Gaps 1;

Qy 184 CCGCTCGCGCCGATGGTCCGGCGCTCATCTGCAGCTGGGAAGTGTGGGATGCGA 243
Db 6 SSRYTSSKRYGKYSSMYSSRSMKYAWGRKYGTSGRGGSRGMMCKWGYRYSY 65
Qy 244 AACAGCAAGCGCGCTCGTGGCGTTCGGCTCCCAAGCGCGCTTGGAGGATCAT 303
Db 66 WGYKWSKSKMKYSGHGTSSKSKRYRYSKRTTCKYRWSKSKWMMKRRKMY 124
Qy 304 CCCATCGTGTGATTCGATTCACCTTCGATTCGATTCGATTCGATTCATTT 363
Db 125 YRMKCYSCASYSYRRCRYTGTMTGMYCKRMCKSKSTRYMTRYTRYWMTGACYGS 184
Qy 364 GTTGGGAGTACCAAAAAGTTTCAGATACGGCTCAAGAGCTCACCCTGACAGCCA 423
Db 185 KMSCKGRSKYSGYKRYGKYTCTSKYKSKMYKSKMYKSKMYKSKMYKSKMYKSKMY 244
Qy 424 GTCATCGTGTGATTCATTAACATTCATTCATTCATTCATTCATTCATTCATTC 483
Db 245 KKSYYTCKRSYYTYRTSTSKGWTGKSKRWSYTWGSKSKRWSYTWGSKSKRWSYTW 304
Qy 484 TGAAGAACCAAAATACAATATACAGGACCAACAATAGTGTTCGATTCAGGAGCAACA 543
Db 305 MSWCMCAWMAWMSRMAWMSRMAWMSRMAWMSRMAWMSRMAWMSRMAWMSRMAWMSR 364
```

```
Db 1783 TCAGGACATGGATTTCTCTCCCAACAAATGTGCAGACTTGGACTTCAAGCTTGTGATATT 1842
QY 394 AGGCTCAAG 403
Db 1843 TCGGTTCAAG 1852

RESULT 10
US-10-101-464A-268
; Sequence 268, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-268

Query Match 3.9%; Score 38.4; DB 4; Length 1107;
Best Local Similarity 49.6%; Pred. No. 0.12;
Matches 127; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 123 ACCACGACGACTCCACGCGTTGAGGGGCTCGCTGAGAACCTAAGCGGCAAAAGGACCG 182
Db 359 AGCTCACGCCCTCATGGCACTCAAGCGCGCCTCGACCCCTCCGCGGGGTCTCACCCT 418
QY 183 TCCGCTCCGCGCGCATGGTCCGCGCCTCATGCTGCAGCTGGGAAGGTGTGGATGCG 242
Db 419 CGTGGTCCGCGCGCGCCGACCTCGCGCGCGCGCTCGTTCGAGGGGGTGGCGTGG 478
QY 243 AAACAGCAAGCGCGCGCTCGTGGCGTCCCGCTCATGCTGCAGCTGGGAAGGTGTGGAGGATCA 302
Db 479 A---CGAGCGCGCGCGCTGGTGAACGCTCGCTGCAGGGGAAGGCGCTTGAGGGGGCGGA 535
QY 303 TCCCATCGTCGATGGTGGAGCTTGATCACCTTCGCTATTGTGATCTCTCGGGTAATTCAT 362
Db 536 TCCGCGCGAGATCGCGCGGCTCGGGAGCTTACCGGGCTGACCTTGACATTCACAGCCCC 595
QY 363 TGGTTGGGGAGGTACC 378
Db 596 TGGCGCGCGAGGTGCC 611

RESULT 11
US-10-101-464A-865
; Sequence 865, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
```

```
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 865
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-865

Query Match 3.9%; Score 38.4; DB 4; Length 2735;
Best Local Similarity 49.6%; Pred. No. 0.21;
Matches 127; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 123 ACCACGACGACTCCACGCGTTGAGGGGCTCGCTGAGAACCTAAGCGGCAAAAGGACCG 182
Db 343 AGCTCACGCCCTCATGGCACTCAAGCGCGCCTCGACCCCTCCGCGGGGTCTCACCCT 402
QY 183 TCCGCTCCGCGCGCATGGTCCGCGGCTCATGCTGCAGCTGGGAAGGTGTGGATGCG 242
Db 403 CGTGGTCCGCGCGCGCCGACCTCGCGCGCGCGCTCGTTCGAGGGGGTGGCGTGG 462
QY 243 AAACAGCAAGCGCGCGCTCGTGGCGTTCGCGGCTCCCAAGCGCGGCTTCGAGGGATCA 302
Db 463 A---CGAGCGCGCGCGCTGGTGAACGCTCGCTCAGGGGAAGGCGCTTGAGGGGGCGGA 519
QY 303 TCCCATCGTCGATGGTGGAGCTTGATCACCTTCGCTATTGTGATCTCTCGGGTAATTCAT 362
Db 520 TCCGCGCGAGATCGCGGGCTCGGGAGCTGACCGGGCTGACCTGCACCTTCAACGCC 579
QY 363 TGGTTGGGGAGGTACC 378
Db 580 TGGCGCGCGAGGTGCC 595

RESULT 12
US-09-228-986-39
; Sequence 39, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-39

Query Match 3.9%; Score 38.2; DB 3; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 217 CTGAGCTGGGAAGGTGTGGGATGCGAATCGAAGCGGCGGCTCGCTGCGGCTTGGCGGT 276
Db 237 CTGCGCTGTGGAACGGAATTCGTGCGAGGCCAAACAGAACGAGTGGTTCATTTCTCT 296
QY 277 CCCCAGCGCGGCGCTTGGAGGGATCATCCCATCGTCGATTGCTGAGCTTGATCACCTTCG 336
```

```

; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(735)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1028)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (424)..(693)
; OTHER INFORMATION: Heat Shock Domain
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(328)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (329)..(1028)
; FEATURE:
; NAME/KEY: Poly A site
; LOCATION: (1028)..(1028)
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: (1)..(1028)
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1028)
; OTHER INFORMATION: Zea Mays L., Line B73
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: (79)..(213)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (736)..(1028)
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(78)
; PUBLICATION INFORMATION:
; AUTHORS: Lund, Adrian A.
; AUTHORS: Blum, Paul H.
; AUTHORS: Bhatramakki, Dinakar
; AUTHORS: Elthon, Thomas E.
; TITLE: Heat-Stress Response of Maize Mitochondria
; JOURNAL: Plant Physiol.
; VOLUME: 116
; PAGES: 1097-1110
; DATE: 1998-03-00
; US-09-249-180-1

```

```

Query Match
Best Local Similarity 4.0%; Score 38.8; DB 3; Length 1028;
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 175 AGGAGCGCTCCGCGCTCCGCGCGCATGCTCCGCGCGCTCATGTCGAGCTGGGAAGGTGT 234
Db 189 AGTAGCGCGCGCTCCGCGGTACACACCGCGCTCCGCTCCGAGCTACGAGGGGC 248
QY 235 GGGATCGGAACAGCAGCGCGCGCTCGTGGGTTGGGCTCCCAAGCGCGCTTGG 294
Db 249 CGAGTCGGAAGACGATAGCTCCGCGAGTACGATGGCGCGCGCGGCGGACTACGC 308
QY 295 AGGGATCATCCCATCTCGATGCTGAGCTGATCACCCTCG 336
Db 309 TGTGCCCGCGCTGTTCTCAGATATTTTCGTGATCGCTTAG 350

```

```

RESULT 8
US-09-434-288-12/c
; Sequence 12, Application US/09434288
; Patent No. 6303767
; GENERAL INFORMATION:
; APPLICANT: Betlach C., Melanie
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA

```

```

; TITLE OF INVENTION: CONSTRUCTS THEREFOR
; FILE REFERENCE: 30062-20030.00
; CURRENT APPLICATION NUMBER: US/09/434,288
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/107,093
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Streptomyces narbonensis
; US-09-434-288-12

```

```

Query Match
Best Local Similarity 4.0%; Score 38.8; DB 3; Length 1476;
Matches 97; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 85 CATCTTGCGAGGTGGCAGGCAACGTCGTGCTGTCGCCACCAACGACCTCCACGCGTT 144
Db 1281 CCTGCCCGCGTGGAGCGGACCGCGCGCTGCGGTTCGTCCACGGGACATCCGCGACGC 1222
QY 145 GAGGGGCTCGCTGAGAACCTTAAGCGCAAGAGAGCGCTCCGCGCTCCGCGCGCATGGTC 204
Db 1221 CGACCTCTCGCGCGGAACTCGCGCGGTGCGACGCGCTCGTCCACTTCGCGCGCGAGAG 1162
QY 205 CGGCGCTCATGCTGACGCTGGAAAGTGTGGATGCCAAACAGCAAGCGCGCGCTCGT 264
Db 1161 CCACGTCGACCGCTCCATCGCGGGGGCGTCCGTGTTCCCGGACCAACGTCACGGGCAC 1102
QY 265 GCGGTTGCGGCTCC 278
Db 1101 CGAGACCTGCTCC 1088

```

```

RESULT 9
US-10-101-464A-840
; Sequence 840, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 840
; LENGTH: 2900
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-10-101-464A-840

```

```

Query Match
Best Local Similarity 4.0%; Score 38.8; DB 4; Length 2900;
Matches 73; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 274 GCTCCCCAAGCGCGCGCTTGGAGGATCATCCCATCGTCGATGTTGAGCTTATCACCCT 333
Db 1723 GCTTTCTTACACGCACTTGAAGGTCTGATACAGCTTCAATCAGTGCAGCTTCAAAATCT 1782
QY 334 TCGCTATTGGATCTCTCGGGTAATTCTGTTGGGAGGTACCAAAAAGTTTTCAGAT 393

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2005, 17:13:16 ; Search time 210 Seconds  
(without alignments)  
7635.958 Million cell updates/sec

Title: US-10-657-852B-3  
Perfect score: 980  
Sequence: 1 gcttcattccatcaaggt.....ccctttacataaaaaaaaaa 980

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.8	4.6	3453	4 US-10-101-464A-861	Sequence 861, App
2	44.8	4.6	7218	1 US-08-232-463-14	Sequence 14, Appl
3	41	4.2	408	4 US-10-101-464A-212	Sequence 212, App
4	40.4	4.1	396	4 US-10-101-464A-221	Sequence 221, App
5	40.4	4.1	3381	3 US-09-336-447A-6	Sequence 6, Appli
6	40.4	4.1	3381	4 US-09-952-267B-6	Sequence 1, Appli
7	38.8	4.0	1028	3 US-09-249-180-1	Sequence 12, Appl
8	38.8	4.0	1476	3 US-09-434-288-12	Sequence 840, App
9	38.8	4.0	2900	4 US-10-101-464A-840	Sequence 268, App
10	38.4	3.9	1107	4 US-10-101-464A-268	Sequence 865, App
11	38.4	3.9	2735	4 US-10-101-464A-865	Sequence 39, Appl
12	38.2	3.9	454	3 US-09-228-986-39	Sequence 39, Appl
13	38.2	3.9	454	4 US-10-101-464A-39	Sequence 421, App
14	38.2	3.9	498	4 US-10-101-464A-421	Sequence 15639, A
15	38.2	3.9	505	4 US-09-621-976-15639	Sequence 2600, Ap
16	38.2	3.9	1590	4 US-09-248-796A-2600	Sequence 837, App
17	38.2	3.9	2691	4 US-10-101-464A-837	Sequence 171, App
18	38.2	3.9	23210	4 US-09-596-002-17	Sequence 166, App
19	38	3.9	535	3 US-09-056-556-171	Sequence 171, App
20	38	3.9	535	3 US-09-072-596-166	Sequence 2, Appli
21	38	3.9	535	4 US-09-072-967-171	Sequence 1, Appli
22	38	3.9	4403765	3 US-09-103-840A-2	Sequence 34, Appl
23	38	3.9	4411529	3 US-09-103-840A-1	Sequence 5, Appli
24	37.4	3.8	4689	3 US-09-105-537-34	Sequence 19, Appl
25	37.4	3.8	36778	3 US-09-105-537-5	Sequence 1, Appli
26	37.4	3.8	38506	3 US-09-320-878-19	Sequence 4, Appli
27	37.4	3.8	38506	4 US-09-141-908-1	

ALIGNMENTS

RESULT 1

US-10-101-464A-861  
; Sequence 861, Application US/10101464A  
; Patent No. 6768041  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; PRIOR FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 861  
; LENGTH: 3453  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-10-101-464A-861  
  
Query Match 4.6%; Score 44.8; DB 4; Length 3453;  
Best Local Similarity 48.4%; Pred. No. 0.002;  
Matches 124; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
  
QY 51 CAAAGTGTTCATGCTGCTGCTTCTTGGGGTTCACTTGCAGGTGGCAGGAGCAACGT 110  
Db 3 CAATGTTCTTTCGC 62  
QY 111 CGTGTGTCGTCGCCACGACGACCTCCAGCGTTGAGGGGCTCGCTGAGAACCTAAGCG 170  
Db 63 CGCGCGTGTCTCTCAACCGAGGAGGCTCTTACCTCCACGCGCGGCGCGGCTCTCCG 122  
QY 171 GCAAGAGGAGCGTCCGCTCTCGCGCGCATGTGTCGCGCGCTCATGCTGAGTGGGAG 230  
Db 123 ACCAGACTCG 182  
QY 231 GTGTGGGATGCGAAACAGCAAGCGCGCTGCTGCGGTTCGCGCTCCCAAGCGGCGCC 290  
Db 183 GCGTCGGGTGCGACCGCGCTCTCGGCTCGCTCTCGTTCGCTCGACCTCTCTCGGCGCGCAACC 242  
QY 291 TTGAGGAGGATCATCCC 306



**This Page Blank (uspto)**

Db 74 GAATGATT 66

Search completed: October 8, 2005, 19:58:39  
Job time : 3177 secs



Db 326 TCCGGGAACAACAATGTGGCTGGGAGCAACAACACTATCGTAAACCGGGAACGACAAT 385  
Qy 632 ACTGTAGTTGGTAGCAACATGTCGTATCAGGACAAGCATATTTGTTACTGATACAAT 691  
Db 386 ACCGTAACTGGTAGCAACATGTCGTATCTGGGGACAACAACATATCGTAACTGACACAAC 445  
Qy 692 AATGTTGTATCCGGGAACGACAATAATGTCTCTGGAAGCTTCCATCTGTATCAGGGGAG 751  
Db 446 AATGCCGTATCCGGGAATGACAATAATGTATCTGGGAGTTTCCATCCGTATCCGGAAC 505  
Qy 752 CACAATACCGTATCCGGGAGCAACAATACTGTATCCGGGAGCAACCATATCGTATCTGGG 811  
Db 506 CACAATACTGTATCTGGGACCAACAACACTGTATCTGGGACAACAACCATGTCGTATCTGGG 565  
Qy 812 AGCAACAAGTCTGTAACAGATGTTAAAT 841  
Db 566 AGCAACAAGTCTGATGAGATGATGATTT 595

## RESULT 13

BJ458121/c

## LOCUS

DEFINITION BJ458121 K. Sato unpublished cdna library, cv. Akashinriki  
vegetative stage leaves Hordeum vulgare subsp. vulgare cdna clone  
baak30m05 3', mRNA sequence.

## ACCESSION

BJ458121

BJ458121.1

GI:21136659

Hordeum vulgare subsp. vulgare

EST.

## SOURCE

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

## FEATURES

Location/Qualifiers

1..689

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Akashinriki"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/clone="baak30m05"

/tissue\_type="leaves"

/dev\_stage="vegetative stage"

/clone\_lib="K. Sato unpublished cdna library, cv.

Akashinriki vegetative stage leaves"

## ORIGIN

Query Match

Best Local Similarity 36.4%; Score 357.2; DB 4; Length 689;

Matches 451; Conservative 0; Mismatches 113; Indels 6; Gaps 2;

Qy 275 CTCCTCAACGCGGCGCTTGGAGGATCATCCCATCGTCGATTGGTGAGCTTGATCACCTT 334

Db 681 CTTGCCAACACAGACTGGTTGGCACCATCCCGTCATGATTGGCGAGCTTGACCACTT 622

Qy 335 CGCTATTGGATCTCTCGGGTAATTCATTGGTTGGGAGGTACCAAAAGTTTGCAGATA 394

Db 621 TACTACTTGGATCTTTCCGATAATTCATTGGTTGGCGAGGTACCCAAGAGTTT---GATA 565

Qy 395 CGGCTCAAGAGGCTCACCACTGACAGCAGTCTACTCGGTATCGGTTCCATTAACATGCTA 454

Db 564 CGGCTCAAGGGCTTCCCATCGTGGTCTTCATCAGGTATGATTTTACTTAACATGCCA 505

Qy 455 TTGCATGT---GAGCAGTAGAAGACGCTCGATGAGAACCAAAATACAATATCAGGAGC 511

Db 504 TTGTATGTGGAGCCTTAACAGAGAATGCTCGACGAACAACCAATACAATATCTGGGAGC 445

Qy 512 AACAAATAGTGTGGATCAGGAGCAACAATGTTGTTCCGGGAATGACAACAACGGTCTGTA 571

Db 444 AACACACTGTAGATCTGGGAGCACAATGTTGTTCTGGGACGACACAACACTGTGCTATA 385

Qy 572 TCTGGGAATAACAACAATGTTCTGGGAGCAACAACAACACTGTTGTTAACTGGAAGTGAAT 631

Db 384 TCCGGGAACAACAACAATGTTGCTGGGAGCAACAACAACACTATCGTAAACCGGAACGACAAT 325

Qy 632 ACTGTAGTTGGTAGCAACCATGTCGTATCAGGAGCAACAACAATGTTTACTGATACAAT 691

Db 324 ACCGTAACCTGGTAGCAACCATGTCGTATCTGGGGACAACAACATATCGTAACTGACAACAAC 265

Qy 692 AATGTTGTATCCGGGAACGACAATAATGTTCTGGAAGCTTCCATCTGTATCAGGGGAG 751

Db 264 AATGCCGTATCCGGGAATGACAATAATGTTATCTGGGAGTTTCCATACCGGTATCCGGAAGC 205

Qy 752 CACAATACCGTATCCGGGAGCAACAATATCTGTATCCGGGAGCAACCATATCGTATCTGGG 811

Db 204 CACAATACTGTATCTGGGAGCAACAACAACACTGTATCTGGGACAACAACCATCTGTATCTGGG 145

Qy 812 AGCAACAAGTCTGTAACAGATGTTAAAT 841

Db 144 AGCAACAAGTCTGATGAGATGATGATTT 115

BF474043 692 bp mRNA linear EST 04-DEC-2000

WHE0840 D01 H02Z5 Wheat vernalized crown cdna library Triticum

aestivum cdna clone WHE0840\_D01\_H02, mRNA sequence.

BF474043

BF474043.1 GI:11543225

EST.

Triticum aestivum (bread wheat)

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

1 (bases 1 to 692)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,

Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,

Seaton, C.L. and Tong, J.C.

The structure and function of the expressed portion of the wheat

genomes - Vernalized crown cdna library

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

Location/Qualifiers

1..692

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE0840\_D01\_H02"

/tissue\_type="Crown tissue of seedling"

/dev\_stage="Five-week old seedling"

/lab\_host="E. coli SOLR"

/clone\_lib="Wheat vernalized crown cdna library"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

## FEATURES

source

```

LOCUS      BJ4622238                684 bp    mRNA    linear    EST 23-MAY-2002
DEFINITION BJ4622238 K. Sato unpublished cDNA library, cv. Akashinriki
            vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
            baak24g16 3', mRNA sequence.
ACCESSION  BJ4622238
VERSION    BJ4622238.1 GI:21140745
KEYWORDS   EST.
SOURCE     Hordeum vulgare subsp. vulgare
ORGANISM   Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 684)
AUTHORS   Sato, K., Saisho, D. and Takeda, K.
TITLE     Barley EST sequencing project in NIG and Okayama Univ
JOURNAL   Unpublished (2002)
COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tschini@genes.nig.ac.jp.
            Location/Qualifiers
FEATURES   source
            1..684
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="Akashinriki"
            /sub_species="vulgare"
            /db_xref="taxon:112509"
            /clone="baak24g16"
            /tissue_type="leaves"
            /dev_stage="vegetative stage"
            /clone_lib="K. Sato unpublished cDNA library, cv.
            Akashinriki vegetative stage leaves"
ORIGIN
Query Match      36.5%; Score 357.6; DB 4; Length 684;
Best Local Similarity 78.4%; Pred. No. 1.5e-95;
Matches 455; Conservative 0; Mismatches 119; Indels 6; Gaps 2;

QY 265 GCGTTGGCGCTCCCAAGCGGCTTCGAGGATCATCCATCGTCGATTGGTGAGCT 324
DB 679 GGAGCTCAACCTTGCACCAACACAGACTGGTTGGCACCATCCCGTATGATTTGGCGAGCT 620
QY 325 TGATCACCTTCGCTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAAG 384
DB 619 TGACCACTTTACTACTTGGATCTTTTCGGATAATTCATTGGTTGGCGAGGTACCAAGAG 560
QY 385 TTTCGAGATACGGCTCAAGAGCTCACCACTGACAGCCAGTCACTCGGTATGGTTCCAT 444
DB 559 TTT---GATACGGCTCAAGGGCTTCGCCATCGCTGCTGCTTCATCAGGTATGATTTTAC 503
QY 445 TAACATGCTATTGCATGT---GAGCAGTAGAAGACGCTCGATGAAGAACCAATACAAAT 501
DB 502 TTAACATGCTATTGTATGTGGAGCCTTAACAGAGATGCTCGAGCAACCAATACAAAT 443
QY 502 ATCAGGGACCAACAATAGTGTGGATCAGGGAGCAACAATGTTGTTTCGGGAATGACAA 561
DB 442 ATCTGGGAGCAACAACACTGTGAGTCTGGAGCACCATTGTTGTTTCGGGAACGACAA 383
QY 562 CAGGTCGTATCTGGGAATAACCACTGTGTCTGGGAGCAACAACACTGTTGTAACCTG 621
DB 382 CACTGTCTATATCGGGAAACAACAACAAATGTGGCTGGGAGCAACAACACTATCGTAA 323
QY 622 AAGTGACATCTAGTGTGGTAGCAACCATGTCGTATCAGGAGCAACAACATATTCTTAC 681
DB 322 GAACGACATATACCGTAATCTGGTAGCAACCATGTCGTATCTGGGGAACAACATATCGTAA 263
QY 682 TGATAACAATAATGTTGTATCCGGGAACGACAAATAATGTCTGGAGCTTCCATCTGT 741
DB 262 TGACAACAATAATGCGGTATCCGGGAATGACATATGATCTCTGGGAGTTCCATACCT 203

```

```

QY 742 ATCAGGGAGCACAAATACCGTATCCGGAGCAACAATACTGTATCCGGAGCAACCATAT 801
DB 202 ATCCGAGAGCCAAATATCTGTATCTGGGACCAACAACACTGTATCTGGGAACAACATGT 143
QY 802 CGTATCTGGGAGCAACAAGTCGTTAACAGATGGTTAATAT 841
DB 142 CGTATCTGGGAGCAACAAGTCGTTAGGAGATGCATGATTT 103

RESULT 12
BJ450137
LOCUS    BJ450137
DEFINITION
            Hordeum vulgare subsp. vulgare
            vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
            baak28d17 5', mRNA sequence.
ACCESSION BJ450137
VERSION    BJ450137.1 GI:21128743
KEYWORDS   EST.
SOURCE     Hordeum vulgare subsp. vulgare
ORGANISM   Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 609)
AUTHORS   Sato, K., Saisho, D. and Takeda, K.
TITLE     Barley EST sequencing project in NIG and Okayama Univ
JOURNAL   Unpublished (2002)
COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tschini@genes.nig.ac.jp.
            Location/Qualifiers
FEATURES   source
            1..609
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="Akashinriki"
            /sub_species="vulgare"
            /db_xref="taxon:112509"
            /clone="baak28d17"
            /tissue_type="leaves"
            /dev_stage="vegetative stage"
            /clone_lib="K. Sato unpublished cDNA library, cv.
            Akashinriki vegetative stage leaves"
ORIGIN
Query Match      36.4%; Score 357.2; DB 4; Length 609;
Best Local Similarity 79.1%; Pred. No. 1.9e-95;
Matches 451; Conservative 0; Mismatches 113; Indels 6; Gaps 2;

QY 275 CTCCCCAAGCGCGCTTGGAGGATCATCCCATCGTCGATTGGTGAGCTTGCATCACCTT 334
DB 29 CTTCGCAACAACAGACTGGTTGGCACCATCCCGTCATGATTTGGCGAGCTTGACCACCTT 88
QY 335 CGCTATTGGATCTCTCGGTAATTCATTGGTTGGGAGGTACCAAAAGTTTGCAGATA 394
DB 89 TACTACTTGGATCTTTCGGATTAATTCATTGGTTGGCGAGGTACCAAGAGTTT---GATA 145
QY 395 CGGCTCAAGAGCCTCACCACTGACAGCCAGTCACCTCGGTATGGTTCATTAAACATGCTA 454
DB 146 CGGCTCAAGGGCTTTCGCCATCGCTGTCATCAGGTATGATTTTACTTAACATGCCA 205
QY 455 TTGCATGT---GAGCAGTAGAAGACGCTCGATGAAGAACCAAAATACAAATATCAGGGACC 511
DB 206 TTGTATGTGGAGCCCTAAACAGAGAATGCTCGAGCAACAACCAAAATACAAATATCTGGGAGC 265
QY 512 AACAAATAGTGTGGATCAGGAGGACCAATGTTGTTTCGGGAATGACAAACCGTTCGTA 571
DB 266 AACAAACTGTGAGATCTGGGAGCACCATGTTGTTTCGGGAACGACAACTGTCATA 325
QY 572 TCTGGGAATAACAACCATGTGCTGGGAGCAACAACACTGTTGTTTAAGTGAAGTGAACAT 631

```



QY 812 AGCAAAAGTCGTAAACAGATGGTTAAATAT-----TCTGTAGGTGACGGA 856  
 Db 146 AGCAAAAGTCGTAGGAGATGATGATTTGTAAGTGGAGTCTCCATCTTCGTGACGGA 87  
 QY 857 TTGCTTCATCTTCCCAAGTTCAGTGTAGCTTACATCAATAGATGGAGACAATCAGTT 916  
 Db 86 GCTACCTTGTGTCCGAGTTCGGTGTAGCTCACAAATCACTTGTGGGGCCAAATCGTGT 27  
 QY 917 ATGTAACCT 924  
 Db 26 ATGAACCT 19

RESULT 8  
 BE490074  
 LOCUS  
 DEFINITION  
 Triticum aestivum cDNA clone WHE0365\_G07\_N13, mRNA sequence.

ACCESSION  
 BE490074  
 SOURCE  
 EST.  
 ORGANISM  
 Triticum aestivum (bread wheat)

REFERENCE  
 AUTHORS  
 Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.

TITLE  
 JOURNAL  
 COMMENT  
 The structure and function of the expressed portion of the wheat genomes - Cold-stressed seedling cDNA library  
 Unpublished (2000)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105959773  
 Fax: 5105959818  
 Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
 Seq primer: StrataGene SK primer.  
 Location/Qualifiers  
 1..588  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE0365\_G07\_N13"  
 /tissue\_type="Seedling"  
 /dev\_stage="Five-day old seedling"  
 /lab\_host="E. coli SOLR"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Five-day old seedlings were transferred to 5 C cold room and kept for 48 hr. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

FEATURES  
 source

ORIGIN  
 Query Match 37.0%; Score 362.4; DB 2; Length 588;  
 Best Local Similarity 79.5%; Pred. No. 5.2e-97;  
 Matches 455; Conservative 0; Mismatches 111; Indels 6; Gaps 2;

QY 257 CGCGTCGTGGCGTTCGGCTCCCAAGCGCGCGCTTGGAGGATCATCCCATCTGTCGATT 316  
 Db 20 CGGCTAGAGGAGCTCAACCTTTGCCAAACAACTGGTCGCTACCATCCCATCTGTCGATT 79  
 QY 317 GGTGAGCTTGATCACTTCGCTATTTGGATCTCTCGGCTAATTCATTCTGTTGGGAGGTA 376  
 Db 80 GGCAGCTTGATCACTTCGCTATTTGGATCTCTCGGCTAATTCATTCTGTTGGGAGGTA 139  
 QY 377 CCAAAAAGTTTGCAGATACGGCTCAAGAGCCTCACCTGACAGCCAGTCTCGGTATG 436  
 Db 140 CCAAGAGTTT--GATACGGCTCAAGGCTCTCGTCATCGTTGGTCTGTTCTACTAGTATG 196  
 QY 437 GGTCCATTAAACATGCTATTTGATGTGA---GCAGTAGAAGAAACCTCGATGAGAACCA 493  
 Db 197 GTTTTACCAACATGCCATTCTATGTGAAGCGTAAATAGAAGAACACTCGACGAAACA 256  
 QY 494 AATCAATATCAGGGACCAACAATAGTCTTGGATCAGGGAGCAACAATGTTGTTTCGGG 553  
 Db 257 AATCAATATCTGGGAGCAACAACACTGTAGATCTGGAAGCACCACCGTTGTTTCTGG 316  
 QY 554 AATGACAAACACGGTCTGATCTGGGAATAAACACCATGTCTTGGGAGCAACACACTGTT 613  
 Db 317 AATGACAAACACTGTCTATCTCGGGAATAAACCAATGTGGCTGTAGCAACACTGTC 376  
 QY 614 GTAACTGGAAGTGACAATCTAGTTAGTGTAGCAACCATGTCTGATATCAGGGAGCAACAGCAT 673  
 Db 377 GTAACTGGAAGTGACAATCTAGTTAGTGTAGCAACCATATTTGATCTCGGGAGCAACAT 436  
 QY 674 ATTGTTACTGATAACAATAATGTTGTATCCGGGAGCAACAATGTTCTCTGGAAGCTTC 733  
 Db 437 ATCGTGACTGACAATAACAATGCCGTATCCGGGAATGACAATAATGTTATCTGGGAGCTTC 496  
 QY 734 CATATCTGTATCAGGGAGCAACAATACCTGTATCCGGGAGCAACAATCTGTTATCCGGAGC 793  
 Db 497 CATACCGTATCCGGGAGCAACAATCTGTTATCTGGGAGCAACAACACTGTATCCGGAAGC 556  
 QY 794 AACCATATCGTATCTGGGAGCAACAACAGTCGT 825  
 Db 557 AACCATGTCTGTTATCTGGGAGCAACAACAGTCGT 588

RESULT 9  
 BE453251  
 LOCUS  
 DEFINITION  
 BJA53251 K. Sato unpublished cDNA library, cv. Akashinriki  
 vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone  
 baak42804 5', mRNA sequence.

ACCESSION  
 BE453251  
 VERSION  
 BE453251.1  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 Hordeum vulgare subsp. vulgare  
 Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Hordeum.

REFERENCE  
 1 (bases 1 to 632)  
 Sato, K., Saiho, D. and Takeda, K.  
 Barley EST sequencing project in NIG and Okayama Univ  
 Unpublished (2002)  
 Contact: Tadao Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.  
 Location/Qualifiers  
 1..632  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Akashinriki"  
 /sub\_species="vulgare"  
 /db\_xref="taxon:112509"

FEATURES  
 source







SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
1 (bases 1 to 686)

REFERENCE Sato, K., Saisho, D. and Takeda, K.  
AUTHORS Barley EST sequencing project in NIG and Okayama Univ  
TITLE Unpublished (2002)  
JOURNAL  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source  
1..686  
Location/Qualifiers

/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mrna"  
/cultivar="Akashinriki"  
/sub\_species="vulgare"  
/db\_xref="taxon:112509"  
/clone="baak46120"  
/tissue\_type="leaves"  
/dev\_stage="vegetative stage"  
/clone\_lib="K. Sato unpublished cdna library, cv.  
Akashinriki vegetative stage leaves"

## ORIGIN

Query Match 37.8%; Score 370.2; DB 4; Length 686;  
Best Local Similarity 75.6%; Pred. No. 2.5e-99;  
Matches 507; Conservative 0; Mismatches 143; Indels 21; Gaps 3;

QY 275 CTCCCCAAGCGGGCTTCGAGGATCATCCCATCGTCGATGGTGAGCTTGATCACCTT 334  
DB 684 TTGTCACAAACAGACTGGTTGGCACCATCCCGTCATGGATTGGCGAGCTTGACCACTT 625  
QY 335 CGCTATTGGATCTCTCGGGTAATTCATTGGTTGGGAGGTACCAAAAGTTTGCAGATA 394  
DB 624 TACTACTTGGATCTTTCCGGATAATTCATTGGTTGGCGAGGTACCCCAAGAGTTT---GATA 568  
QY 395 CGSCTCAAGAGCTCACCACTGCACGCCAGTCACTCGGTATGGTTCCATTACATGCTA 454  
DB 567 CGGCTCAAGGGCTTCGCCATCGCTGTCGTTCACTAGGATAGATTTTACTACATGCCA 508  
QY 455 TTGTCATGT---GAGCAGTAGAAGAACGCTCGATGAAGAACCAATAACAATATCAGGAGC 511  
DB 507 TTGTATGTGGAGCCTAACAGAAGATGCTCGACGAACCAACCAATAACAATATCTGGAGC 448  
QY 512 AACATATGTTGGATCAGGAGCAACAATGTTGTTTCGGGAATGACACAGGTCGTA 571  
DB 447 AACCAACATGTCAGATCTGGAGCACCAATGTTGTTCTGGGAACGACCAACATGTCATA 388  
QY 572 TCTGGGAATAACAACATCTGTCGGAGCAACACACTGTTCTTAACCTGGAAGTGACAT 631  
DB 387 TCCGGAAACAACAACATGTGGCTGGAGCAACACACTATCGTAACCGGGAAACGACAAT 328  
QY 632 ACTGTAGTTGGTAGCAACCATGTGCTATCAGGGAACAAAGCATATGTTACTGATACAAT 691  
DB 327 ACCGTAACTGGTAGCAACCATGTGCTATCTGGGACAAACATATCGTAACGACAAAC 268  
QY 692 AATGTTGTATCCGGGAACGACAAATATGTTGTCGGAAGCTTCCATATCTGATCAGGGAG 751  
DB 267 AATGCCGTATCCGGGAATGACAATAATGATATCTGGAGTTTCCATACCGGTATCCGAAGC 208  
QY 752 CACATACCGTATCCGGGAGCAACATCTGTATCCGGAGCAACCAATATCGTATCTGGG 811  
DB 207 CACATACCTGATCTCGGAGCAACCAACACTGTATCTGGGAACCAACCATGTCGTATCTGG 148  
QY 812 AGCAACAAAGTCGTACAGATGGTTAATAT-----TCTGTAGGTGACGGA 856  
DB 147 AGCAACAAAGTCGTAGGAGATGATGATTTGTAAGTGGAGTGTCCATCTTCGTGACGGA 88

QY 857 TTGCTTCCATCTTCCCAAGTTCAGTGTAGCTTAATAGATGAGACATACGTT 916  
DB 87 GCTCACCTTGTTCGAGTTCGGTGTAGCTCACAATCACTTGGTGGGCCAATCGTGT 28  
QY 917 ATGTAACCTCA 927  
DB 27 ATGTAACCTCA 17

## RESULT 4

BJ461803/c  
LOCUS  
DEFINITION BJ461803 K. Sato unpublished cdna library, cv. Akashinriki  
vegetative stage leaves Hordeum vulgare subsp. vulgare cdna clone  
baak46n01 3', mRNA sequence.

ACCESSION BJ461803  
VERSION BJ461803.1 GI:21140313  
SOURCE EST.  
ORGANISM Hordeum vulgare subsp. vulgare  
Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 699)  
AUTHORS Sato, K., Saisho, D. and Takeda, K.  
TITLE Barley EST sequencing project in NIG and Okayama Univ  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source  
1..699  
Location/Qualifiers

/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mrna"  
/cultivar="Akashinriki"  
/sub\_species="vulgare"  
/db\_xref="taxon:112509"  
/clone="baak46n01"  
/tissue\_type="leaves"  
/dev\_stage="vegetative stage"  
/clone\_lib="K. Sato unpublished cdna library, cv.  
Akashinriki vegetative stage leaves"

## ORIGIN

Query Match 37.8%; Score 370.2; DB 4; Length 699;  
Best Local Similarity 75.6%; Pred. No. 2.5e-99;  
Matches 507; Conservative 0; Mismatches 143; Indels 21; Gaps 3;

QY 275 CTCCCCAAGCGGGCTTCGAGGATCATCCCATCGTCGATGGTGAGCTTGATCACCTT 334  
DB 686 TTGTCACAAACAGACTGGTTGGCACCATCCCGTCATGATGCGAGCTTGACCACTT 627  
QY 335 CGCTATTGGATCTCTCGGGTAATTCATTGGTTGGGAGGTACCAAAAAGTTTGCAGATA 394  
DB 626 TACTACTTGGATCTTTCCGGATAATTCATTGGTTGGCGAGGTACCCCAAGAGTTT---GATA 570  
QY 395 CGSCTCAAGAGCTCACCACTGCACGCCAGTCACTCGGTATGGTTCCATTACATGCTA 454  
DB 569 CGGCTCAAGGGCTTCGCCATCGCTGTCGTTCACTCAGGTATGATTTTACTACATGCCA 510  
QY 455 TTGTCATGT---GAGCAGTAGAAGAACGCTCGATGAAGAACCAATAACAATATCAGGAGC 511  
DB 509 TTGTATGTGGAGCCTAACAGAAGATGCTCGACGAACAAACAAATATCTGGAGC 450  
QY 512 AACCAATATGTTGGATCAGGGAGCAACAATGTTGTTTCGGGAATGACCAACGTTGCTA 571  
DB 449 AACCAACTGTCCAGATCTGGGAGCCAAATGTTGTTTCTGGGAACGACCAACTGTCTA 390

```

QY 293 GGAGGATCATCCATCGTCGATGGTGAGCTTGATCACCTCGCTATTGGATCTCTCG 352
Db 87 GTTGGACCATCCCGTCTGGATTGGCGAGCTTGACACCTTTACTTGGATCTTTCG 146
QY 353 GGTAAATTCATTGGTTGGGAGAGTACCAAAAGTTTTCGATACGGCTCAAGAGCCTCAC 412
Db 147 GATAATTCATTGGTTGGCGAGGTACCCAGAGTTT--GATACGGCTCAAGGGCTTCGCC 203
QY 413 ACTGACAGCACTGCTACTCGGTATGGTTCCATTAACATGCTATTGATGT---GAGCAGT 469
Db 204 ATCGTGGTCTGATCATCAGGTATGATTTTACTAATGCTTATGTTGGAGCCTAAC 263
QY 470 AGAAGACGCTCGATCAAGAACCAATAAATAATACAGGACCAACAATAGTCTGATCA 529
Db 264 AGAAGAAATGCTCGAGCAACAACAATAAATAATCTGGGAGCAACAACATCTGATCT 323
QY 530 GGGAGCAACAATGTTGTTTCCGGGAATGACACACGGTCTGATCTGGGAATAAACCAAT 589
Db 324 GGGAGCAACAATGTTGTTTCTGGGAACGACACATCTGTCATATCCGGGAACAACAAT 383
QY 590 GTGCTGGAGCAACAACACTGTTGTAATCGGAAGTGACAATCTGTTAGTTGGTAGCAAC 649
Db 384 GTGGCTGGAGCAACAACACTATCGTAACCGGGAACGACAATACCGTAACTGGTAGCAAC 443
QY 650 CATGCTGATCAGGACAAAGCATATGTTACTGATAACAATAATGTTGTTATCCGGGAAC 709
Db 444 CATGCTGATCAGGACAAACATATCGTAACCTGACACAACAATGCGGTATCCGGGAAT 503
QY 710 GACAATAATGTTCTGGAAGCTTCCATCTGATCAGGGGACCAACAATACCGTATCCGGG 769
Db 504 GACAATAATGTTCTGGAAGTTCCATCTGATCAGGTATCCGGGAACCAATATCTGTTGG 563
QY 770 AGCAACAATCTGATTCGGGAGCAACCATATCGTATCTGGAGCAACAAGTCGTAAACA 829
Db 564 ACCAACAACACTGATCTCGGAACACCATGTCGTATCTGGAGCAACAACAAGTCGTAGGA 623
QY 830 GATGTTAATAT-----TCTGTAGTGCAGGATGCTTCCATCTTCCCA 874
Db 624 GATGATGATTTGTAAGTGGAGTGTCCATCTTCCGTGACGGAGCTCACCTTGTGTCCGA 683
QY 875 GTTCAGTGTAGCTTACAATCAATAGATGGAGCAATCAGCTTATGTAACCTCA 927
Db 684 GTTCGTGTAGCTCACATCACTTGTGGGGCAATCGTGTATGTAACCTCA 736

```

```

RESULT 2
BJ458554/c
LOCUS
DEFINITION
746 bp mRNA linear EST 23-MAY-2002
BJ458554 K. Sato unpublished cDNA library, cv. Akashinriki
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak32h16 3', mRNA sequence.
ACCESSION
BJ458554
VERSION
BJ458554.1 GI:21137090
KEYWORDS
EST.
SOURCE
Hordeum vulgare subsp. vulgare
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 746)
REFERENCE
Sato, K., Saisho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Source
Location/Qualifiers
1..746
/organism="Hordeum vulgare subsp. vulgare"

```

```

/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak32h16"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"
ORIGIN
Query Match 38.0%; Score 372.4; DB 4; Length 746;
Best Local Similarity 73.7%; Pred. No. 5.5e-100;
Matches 523; Conservative 0; Mismatches 166; Indels 21; Gaps 3;
QY 236 GGATCGAAAACAGCAGCGCGCTCGTGGGTTCGGGCTCCCAAGCGCGCTTGA 295
Db 717 GCATCTTGGCAGGCTCGCATGTTGGAGAGCTCAACCTTGGCAACAACAGACTGGTT 658
QY 296 GGGATCATCCCATCGTTCGATTGGTGAGCTTCATACCTTCGCTATTTCGATCTCTCGGT 355
Db 657 GGCACATCCCGTCATGATTGGCGAGCTTGACACCTTTACTTTCGATCTTTCGAT 598
QY 356 AATTCATTGTTGGGAGGTACCAAAAAGTTTTCGAGATACCGCTCAAGAGCCTCACCT 415
Db 597 AATTCATTGTTGGCAGGTACCCAGAGTTT---GATACGGCTCAAGGGCTTCGCCATC 541
QY 416 GACAGCAGTCACTCGGTATGGGTTCCATTAACTGCTATTTCATGT---GAGCAGTAGA 472
Db 540 GCTGCTCGTTCATCAGGTATGATTTTACTAATCATGCTTGTATGTGGAGCCTAACAGA 481
QY 473 AGACGCTCGATGAGAACCAATAATATCAGGAGCAACAATAGTGTGGATCAGGG 532
Db 480 AGAATGCTCGACGAACAACAATAACAATATCTGGAGCAACAACACTGTGATCTGG 421
QY 533 AGCAACAATGTTGTTTCCGGGAATGACAACACGGTCTGATCTGGGAATAACAACCATGTG 592
Db 420 AGCACCATTGTTTCTGGGAACGACAACACTGTTCATATCCGGGAACAACAACATGTG 361
QY 593 TCTGGAGCAACAACACTGTTGTAATCTGGAAGTGACAATCTGTTAGTTGGTAGCAACCT 652
Db 360 GCTGGAGCAACAACACTATCGTAACCGGGAACGACAATACCGTAACCTGGTAGCAACCT 301
QY 653 GTCGTATCAGGAGCAACAACATATGTTTACTGATAACAATATGTTGTTATCCGGGAACGAC 712
Db 300 GTCGTATCTGGGACAACAATATCGTAACCTGACAACAACAATGCGGTATCCGGGAATGAC 241
QY 713 AATAATGTTCTGGAGCTTCCATCTGATCAGGGGAGCAACAATACCGTATCCGGGAGC 772
Db 240 AATAATGTTCTGGAGGTTTCCATACCGTATCCGGGAACCAACAATCTGATCTGGGACC 181
QY 773 AACAACTACTGTATCCGGGAGCAACAATATCGTATCTGGAGCAACAAGTCTGTACAGAT 832
Db 180 AACAACTACTGTATCTGGGAACAACACTGTCGTATCTGGAGCAACAAGTCTGTAGAGAT 121
QY 833 GGTAAATAT-----TCTGTAGTGCAGGATGCTTCCATCTTCCCAAGTT 877
Db 120 GCATGATTGTAAAGTGGAGTGTCCATCTTCGTGACGGAGCTCACCTTGTGTTGTCGAGTT 61
QY 878 CAGTGTAGCTTACAATCAATAGATGAGACAATACGTTATGTAACCTCA 927
Db 60 CGGTGTAGCTACAATCACTTGGTGGGCAATCGTGTATGTAACCTCA 11

```

```

RESULT 3
BJ461731/c
LOCUS
DEFINITION
686 bp mRNA linear EST 23-MAY-2002
BJ461731 K. Sato unpublished cDNA library, cv. Akashinriki
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak46120 3', mRNA sequence.
ACCESSION
BJ461731
VERSION
BJ461731.1 GI:21140241
KEYWORDS
EST.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2005, 13:12:20 ; Search time 3171 Seconds  
(without alignments)  
11763.792 Million cell updates/sec

Title: US-10-657-852B-3  
Perfect score: 980  
Sequence: 1 gcttgattccatcaaggt.....cccttacataaaaaaaaaa 980

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues 68479088

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: gb\_est1:\*
  - 2: gb\_est2:\*
  - 3: gb\_hic:\*
  - 4: gb\_est3:\*
  - 5: gb\_est4:\*
  - 6: gb\_est5:\*
  - 7: gb\_est6:\*
  - 8: gb\_gss1:\*
  - 9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372.8	38.0	767	4	BJ451048
2	372.4	38.0	746	4	BJ458554
3	370.2	37.8	686	4	BJ461731
4	370.2	37.8	699	4	BJ461803
5	366.6	37.4	687	4	BJ459099
6	365.6	37.3	693	4	BJ457564
7	365.6	37.3	696	4	BJ461908
8	362.4	37.0	588	2	BE490074
9	362.2	37.0	632	4	BJ453251
10	359.4	36.7	671	4	BJ454271
11	357.6	36.5	684	4	BJ452238
12	357.2	36.4	609	4	BJ450137
13	357.2	36.4	689	4	BJ458121
14	352	35.9	692	2	BF474043
15	348.2	35.5	703	2	BE705098
16	347	35.4	625	4	BJ448772
17	347	35.4	726	4	BJ451602
18	347	35.4	830	2	BE705403
19	346.4	35.3	830	2	BE705403
20	344.4	35.1	663	4	BJ461352
21	340.4	34.7	830	2	BE705439
22	335.6	34.2	1115	7	CK214468
23	332.6	33.9	841	4	BJ224369
24	331.6	33.8	851	4	BJ448689

C 25	328.8	33.6	675	4	BJ460799
C 26	325.4	33.2	724	4	BJ456413
C 27	323	33.0	620	4	BJ447068
C 28	322.2	32.9	589	1	AV909087
C 29	318.6	32.5	862	4	BJ454200
C 30	316.2	32.3	818	7	CK197682
C 31	316.2	31.7	775	2	BE705684
C 32	301.6	30.8	709	4	BJ453816
C 33	296.8	30.3	648	4	BJ300903
C 34	296	30.2	574	1	AV910970
C 35	294.6	30.1	914	7	CK156167
C 36	294	30.0	612	1	AV911379
C 37	293.6	30.0	581	4	BM376553
C 38	292.8	29.9	588	4	BJ453528
C 39	285	29.1	480	1	AJ460325
C 40	277.8	28.3	622	1	AU251218
C 41	275.8	28.1	585	4	BI479842
C 42	275	28.1	841	7	CK196896
C 43	263.2	26.9	532	2	BF200590
C 44	255.6	26.1	420	1	AJ460320
C 45	255.6	26.1	420	1	AJ460322

ALIGNMENTS

RESULT 1  
BJ451048  
LOCUS  
DEFINITION  
          BJ451048 K. Sato unpublished cDNA library, cv. Akashinriki  
          vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone  
          baak32h16 5', mRNA sequence.  
ACCESSION  
          BJ451048  
VERSION  
          BJ451048.1 GI:21129647  
KEYWORDS  
          EST.  
SOURCE  
          Hordeum vulgare subsp. vulgare  
ORGANISM  
          Hordeum vulgare subsp. vulgare  
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
          Poideae; Triticeae; Hordeum.  
REFERENCE  
          1 (bases 1 to 767)  
          Sato, K., Saisho, D. and Takeda, K.  
AUTHORS  
          Barley EST sequencing project in NIG and Okayama Univ  
TITLE  
          Unpublished (2002)  
JOURNAL  
          Contact: Tadasu Shin-i  
COMMENT  
          Center For Genetic Resource Information  
          National Institute of Genetics  
          1111 Yata, Mishima, Shizuoka 411-8540, Japan  
          Tel: 81-559-81-6856  
          Fax: 81-559-81-6855  
          Email: tshini@genes.nig.ac.jp.

FEATURES

source  
          1..767  
          /organism="Hordeum vulgare subsp. vulgare"  
          /mol\_type="mRNA"  
          /cultivar="Akashinriki"  
          /sub\_species="vulgare"  
          /db\_xref="taxon:112509"  
          /clone="baak32h16"  
          /tissue\_type="leaves"  
          /dev\_stage="vegetative stage"  
          /clone\_lib="K. Sato unpublished cDNA library, cv.  
          Akashinriki vegetative stage leaves"

ORIGIN

Query Match 38.0%; Score 372.8; DB 4; Length 767;  
Best Local Similarity 73.5%; Pred. No. 4.2e-100;  
Matches 524; Conservative 0; Mismatches 168; Indels 21; Gaps 3;  
QY 233 GTGGATGCGAAACGACGACGCGCGTCTGGTGGCGTTTCGGCTCCCAAGCGCGCCTT 232  
DB 27 GNAGCATCTTGGAGGCTTCCATCTTGGAGGAGCTCAACCTTGCACACACACTG 86

**This Page Blank (uspto)**

CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants  
CC have a high nutritional value with reduced apical dominance or dwarfism,  
CC early flowering or altered metabolic pathways. This sequence represents a  
CC plant nucleic acid of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification but was obtained in  
CC electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX

SQ Sequence 4536 BP; 1102 A; 1247 C; 1221 G; 965 T; 0 U; 1 Other;

Query Match	8.1%;	Score 79.4;	DB 12;	Length 4536;
Best Local Similarity	63.1%;	Pred. No. 2.2e-13;		
Matches 157;	Conservative 0;	Mismatches 86;	Indels 6;	Gaps 2;

  

QY	104	GCAACGTCGTGTCGTGCCACACGACCTCCACGCGTTGAGGGGCGCTCGCTGAGAAC	163
Db	76	GCTGGGGGGCGCATGCCACCCGGAGAGACTCTCGCGCTGCGGGCATTTGCGGGGAAT	135
QY	164	CT---AAGCGGCAAGGAGCGTCCGCTCCGCGCCGCATGTTCCGCGCGCTCATGCTGC	220
Db	136	CTCTCTCGGGTGGGGGGCGCGGACTCCGCGCGCGTGTCTCGGGTGACGCCCTGCTGC	195
QY	221	AGCTGGGAAGGTGTGGATCGGAACAGCAAGCGGGCGCGTCTGTGGCGTTGCGGCTCCCC	280
Db	196	GCCTGGGACGCGCTCGCTGCGACGCGC---CGCCCGAGTCAAGCGCGCTGCGCCTCCCC	252
QY	281	AAGCGGCGCTTGGAGGGATCATCCCATCGTCGATTGGTGAGCTTGATCACCTTCGCTAT	340
Db	253	GGGCGAGGTCTCGAGGGGGCCATCCCGCCCTCCTCGCGCCCTCGCGCGCTCCAGGAC	312
QY	341	TTGGATCTC	349
Db	313	CTCGACCTC	321

Search completed: October 8, 2005, 17:53:50  
Job time : 580 secs

PD 18-MAR-2004.  
 XX 09-SEP-2003; 2003WO-NZ000199.  
 XX 09-SEP-2002; 2002US-0409557P.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (WRIGHT-) WRIGHTSON SEEDS LTD.  
 XX Demmer J, Shenk MA, Hall C, Fish SA;  
 PI WPI; 2004-248453/23.  
 XX P-PSDB; ADM41471.  
 DR New antifreeze proteins and encoding polynucleotides, useful for  
 PT modulating cold tolerance in organisms, as food additives, or for  
 PT treating tumors or disorders associated with the presence of unwanted  
 PT biocrystals (e.g. gout).  
 XX Claim 1; SEQ ID NO 1; 71pp; English.  
 XX The present sequence is that of cDNA encoding APPI, an antifreeze protein  
 CC of perennial ryegrass. The cDNA was isolated from a pseudostem cDNA  
 CC expression library. The invention provides forage grass (perennial  
 CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides  
 CC encoding them ADM41458-ADM41483. The polynucleotides were isolated from  
 CC tissues taken at different times of year (winter and spring) and from  
 CC different parts of the plants. The polynucleotides can be used to  
 CC modulate the cold tolerance of an organism, especially plants, mammals,  
 CC insects, fungi, archaea and bacteria. The method involves incorporating  
 CC an antifreeze polynucleotide, under the control of a gene promoter  
 CC sequence, into the genome of the organism, or introducing double-stranded  
 CC RNA corresponding to the polynucleotide into the cells of the organism,  
 CC thereby inhibiting expression of an antifreeze polypeptide. The  
 CC antifreeze protein can be used for the cryopreservation of a cell or  
 CC tissue, as a food additive of a frozen food product, in a method for  
 CC decreasing the time required to dehydrate a composition, to treat a  
 CC disorder characterised by biocrystals associated with disorders such as  
 CC gout and kidney stones, to preserve the viability of a molecular biology  
 CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
 CC to protect a plant from damage due to frost or freezing.  
 XX Sequence 841 BP; 216 A; 217 C; 222 G; 186 T; 0 U; 0 Other;  
 SQ

Query Match 26.8%; Score 262.2; DB 12; Length 841;  
 Best Local Similarity 73.5%; Pred. No. 3.1e-71;  
 Matches 349; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

QY 380 AAAAGTTTCAGATACGGCTCAAGAGCCTCACCACCTGACAGCCAGTCACTCGGTATGGT 439  
 DB 262 AAGAGCTTCCAGATATTGCTCAAGAGGCTCACCGCGCTGCGGTCTACTGGTAAGCG 321  
 QY 440 TCATTAAATGCTATTGTCATGTGAGCAGTAG---AAGAACGCTCGATGAAGAACCAAT 496  
 DB 322 TTCACTCACATGCCATTATCATGTGAAGCCTAGCCAGGAACACTCGACGAGACCAAT 381  
 QY 497 ACAATATCAGGACCAACAATAGTGTGATCAGGAGCAACAATGTTTTCGGGAT 556  
 DB 382 ACAATAACTGGGATCAACAATACTGTGCATCGGAGCAACAATGTTTTCGGGAC 441  
 QY 557 GACACACGGTCTGATCTCGGGAATACCAACCTGTCGTGGAGCAACAACACTGTGTA 616  
 DB 442 GATAACACTGTATATTCGGGAACAACAACGTCGTGTCGGGAGCCACACACCGTGTGA 501  
 QY 617 ACTGGAAGTGACAATCTAGTGTGTAGCAACCATGTCGTATCAGGCAACAAGCATAT 676  
 DB 502 TTTGGGGTGACAACCTTCAATAGTGAAGTACCATGTCGTATCTGGGAACCAACATGTT 561  
 QY 677 GTTACTGATAACAATAATGTTGTATCCGGGAGCAACAATAATGTTGTCTGGAAGCTTCCAT 736  
 DB 562 GTGACTGACAACAAGAAATCCCGTATCCGGGGAGCAACAATACTGTATCTGGAAGCAAAAT 621  
 QY 737 ACTGTATCAGGAGGAGCACAAATACCGTATCCGGGAGCAACAATACTGTATCTCGGAGCAC 796

DB 622 ACCGTATCCGGAGACCAACCATCGATCTGATCTGGAGCCACAGTACGTATCCGGAGCAC 681  
 QY 797 CATATCGTATCTGGAGCAACAAGTCTGTAACAGATGGTTAATATTCCTGTTAGGTG 851  
 DB 682 AATACGGTATCTGGAGAAACAATTCCTGATATATGGGAACAACAATATTTGTATCTG 736

RESULT 15  
 ADJ40496  
 ID ADJ40496 standard; cDNA; 4536 BP.  
 XX AC ADJ40496;  
 XX DT 06-MAY-2004 (first entry)  
 XX DE Plant cDNA #1496.  
 XX KW Plant; gene; ss; transcription; plant genome augmentation; cereal;  
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;  
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;  
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;  
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;  
 KW antifungal.  
 OS Eukaryota.  
 XX US2004016025-A1.  
 XX 22-JAN-2004.  
 XX 26-SEP-2002; 2002US-00260238.  
 XX 26-SEP-2001; 2001US-0325277P.  
 XX 26-SEP-2001; 2001US-0325448P.  
 XX 04-APR-2002; 2002US-0370620P.  
 XX (BUDW/) BUDWORTH P.  
 PA (MOUG/) MOUGHAMER T.  
 PA (BRIG/) BRIGGS S P.  
 PA (COOP/) COOPER B.  
 PA (GLAZ/) GLAZEBROOK J.  
 PA (GOFF/) GOFF S A.  
 PA (KATA/) KATAGIRI F.  
 PA (KREP/) KREPS J.  
 PA (PROV/) PROVART N.  
 PA (RICK/) RICKE D.  
 PA (ZHUT/) ZHU T.  
 XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;  
 PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;  
 XX WPI; 2004-190374/18.  
 XX New rice promoter, useful for manipulating crop plants to alter or  
 PT improve phenotypic characteristics, e.g. produce large quantities of oil  
 PT or proteins, resistance to insecticides, virus or fungi, stress tolerance  
 PT or high nutritional value.  
 XX Claim 25; SEQ ID NO 1496; 230pp; English.  
 XX The invention relates to plant nucleotide sequences that direct seed-,  
 CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential  
 CC or constitutive transcription of an operatively linked nucleic acid  
 CC segment. The invention also relates to a method for augmenting a plant  
 CC genome and a method of identifying a gene, where its expression is  
 CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive  
 CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,  
 CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,  
 CC sorghum, rice or wheat. The polynucleotides and the polypeptides they  
 CC encode are useful for manipulating crop plants to alter or improve  
 CC phenotypic characteristics, to produce large quantities of oil or  
 CC proteins, to incur resistance to insecticides, viruses or fungi, and to

Qy	383	AGTTTGCAGATACGGCTCAAGAGCCTCACCACTGACAGCCAGCTCACTCGGTATGGGTCC	442
Db	292	AGTTTCCAGATATTCTCTCAAGGGCTCACACCGCTGGCCGTCTCACTGGGTAAAGCGCTTC	351
Qy	443	ATTAAACATGCTATTGCATGTGTGACAGTAG--AAGAACGCTCGATCGAGAACCAANTACA	499
Db	352	ACTAACATGCCATTACATGTGAAGTCTTAGCCAAAGAAACATCGACGAGAAACACAAATACA	411
Qy	500	ATATCAGGGACCAACAAATAGTGTTCGATCAGGGAGCAACAATGTTGTTTCCGGGAATGAC	559
Db	412	ATAACTGGGATCAACAATACTGTCAAAATCCGGGAGCAACAAATGTTGTTTCTGGGAACGAT	471
Qy	560	AACACGGTCTGATCTGGGAATAACCAATGTGTCTGGGAGCAACAACACTGTTGTAACT	619
Db	472	AACACTGTCTATTCGGGAAACAACACGTCGTGTCCGGGAGCCACAACACCGTCGTATTT	531
Qy	620	GGAAGTGAACAATACGTGTAGTTGGTAGCAACCAATGTCTGATCAGGGACAAAGCATATTGTT	679
Db	532	GGGGGTGACAAATTTCTTAAGCGGTAGCAACCAATGTCGATCTCTGGGAAACCAACATGTCGTG	591
Qy	680	ACTGATAACAATAATGTTGTATCCGGGAACGACAATAATGTCTCTGGGAAGCTTCCATACT	739
Db	592	ACTGACAAACAAGATGCCGTATCCGGGAGCCACAATACGTGTATCTGGGAAGCCAAATATCC	651
Qy	740	GTATCAGGGGAGCAATATACCGTATCCGGGAGCAACAATACTGTATCCGGGAGCAACCAAT	799
Db	652	GTATCCGGGAACCAATATCATATCTCGAGCCACAGTACCATATCCGGGAACCACAAT	711
Qy	800	ATCGTATCTGGGAGCAACAAGTCGTAAACAGATGTTTAATATCTGTAGGTG	851
Db	712	ACCGTATCCGGGAGCAACAATTTTGGTATCTGGGAACAACAATTAATTTGTATCTG	763

```

RESULT 13
AAX99717
ID   AAX99717 standard; DNA; 357 BP.
XX
XX
AC   AAX99717;
XX
XX
DT   29-SEP-1999 (first entry)
XX
XX
DE
DE
DE
KW   Grass anti-freeze protein coding sequence.
KW   Anti-freeze protein; grass; plant; frozen food product; frost tolerance;
KW   frozen confectionery; ss.
XX
XX
OS   Lolium perenne.
XX
XX
WO9937782-A2.
XX
XX
PD   29-JUL-1999.
XX
XX
PF   23-DEC-1998; 98WO-EP008553.
XX
XX
PR   22-JAN-1998; 98GB-00001408.
XX
XX
PA   (UNIL ) UNILEVER NV.
PA   (UNIL ) UNILEVER PLC.
XX
XX
PI   Jarman CD, Sidebottom CM, Twigg S, Worrall D;
XX
XX
DR   WPI; 1999-458697/38.
DR   P-PSDB; AAY22472.
XX
XX
PT   New plant anti-freeze protein useful in frozen food products.
XX
XX
PS   Claim 8; Page 37; 39pp; English.
XX
XX
CC   This sequence encodes the plant anti-freeze protein of the invention. The
CC   anti-freeze protein is characterised in that at least 40% of its amino
CC   acids are from the group of serine, threonine and asparagine. The anti-
CC   freeze protein can be used in frozen food products, especially frozen

```

CC	confectionery. Anti-freeze proteins are especially used in food products,									
CC	which are heated, e.g. by pasteurisation, blanching or sterilisation									
CC	prior to freezing. Plants transformed with a nucleic acid sequence									
CC	encoding the anti-freeze protein have an increased frost tolerance. Prior									
CC	art anti-freeze proteins have not been applied to commercially available									
CC	food products, due to high costs and complicated processes for obtaining									
CC	the protein. Also prior art anti-freeze proteins have tended to									
CC	destabilise during processing especially during the pasteurisation step.									
CC	This is overcome by the present anti-freeze protein. The anti-freeze									
CC	proteins provide an ice particle size following an ice recrystallisation									
CC	inhibition assay of 15 mu M or less. The anti-freeze protein ingredient									
CC	means that mixes can be frozen under: quiescent conditions, e.g. in a shop									
CC	or home freezer without the formation of unacceptable ice crystal shapes									
CC	and hence with a texture different to products normally obtained via									
CC	quiescent freezing									
XX										
SQ	Sequence 357 BP; 116 A; 80 C; 89 G; 72 T; 0 U; 0 Other;									
	Query Match	27.0%;	Score 264.2;	DB 2;	Length 357;					
	Best Local Similarity	83.8%;	Pred. No. 4.4e-72;							
	Matches 299;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;					
QY	482	GATCAAGACCAATACATATCAGGGACCAACAATAGTGTGGATCAGGAGCAACAAT	541							
DB	1	GATGAACAGCCGAATACGATTTCTGGGAGCAACAATACTGTTCAGATCCGGGAGCAAAAAT	60							
QY	542	GTGTGTTTCCGGGAATGACAAACCGTTCGTATCTGGGAATAACAAACCATGTGTCTGGGAGC	601							
DB	61	GTTCCTGCTGGGAATGACAACACCGTCATATCTGGGGACAAACAATAGTGTGTCTGGGAGC	120							
QY	602	AACAACACTGTTCTGTAACCTGGAAGTGACAATACATCTGTATGTTGGTGAAGAACCAATGTCGTATCA	661							
DB	121	AACAACACTGTCGTAAGTGGGAATGCAATACCGTAAACCGGACCAACCATGTGTCGTATCA	180							
QY	662	GGGACAAGCAATATTGTTACTGTATACATAATATGTTGTATCCGGGAACGACATAATGTG	721							
DB	181	GGGACAACCAATATCGTTTACAGACAACAACAAATAAACGTATCCGGGAACGATAATAATGTA	240							
QY	722	TCCTGGAGCTTCCCATCTACTGTATCAGGGGAGCAACAATACCGTATCCGGGAGCAACAATACT	781							
DB	241	TCCGGGAGCTTTCATACCGTATCCGGGGGCAACAATACTGTGTCCGGGAGCAACAATAACC	300							
QY	782	GTATCCGGGAGCAACCATATCTGTATCTGGGAGCAACAAGTTCGTAAACAGATGTTAA	838							
DB	301	GTATCTGGGAGCAACCAACGTTGTATCTGGAAGCAACAAGTCGTGACAGACGCTTAA	357							
RESULT 14										
ADM41458	ID	ADM41458 standard; cDNA; 841 BP.								
XX	AC	ADM41458;								
XX	DT	03-JUN-2004 (first entry)								
XX	DE	Perennial ryegrass antifreeze protein AFPI cDNA.								
XX	KW	Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;								
XX	KW	antigout; litholytic; nephrotropic; cytostatic; gene; ss.								
OS	XX	Lolium perenne.								
XX	Key	Location/Qualifiers								
FT	CDS	55..786								
FT		/*tag= b								
FT	sig_peptide	55..114								
FT		/*tag= a								
FT	mat_peptide	115..783								
FT		/*tag= c								
XX	XX									
XX	XX	WO2004022700-A2.								
XX	PN									



Qy	26	GTTCATCCATGCTCGTGAATACATGGCAAAAGTGTTCATGCTGCTGCTTCTTCTTGGGGTTC	85
Db	25	GTTTCTTTGGCATACCAACAATAGCTGAAATCCATGGGGCTGTGTGCTCTTCTTTGGCGTTC	84
Qy	86	ATCTTGAGAGTGGCAGAGCAAACTGCTGTGTGCTGCGCACCAACAGACACCTCCACACGGTTG	145
Db	85	CTCTTGCAGTGGCATGCGCGGCAACGTGCAGTTTGCACCCCGATGACCTCCGTGCGCTG	144
Qy	146	AGGGGCTCGTGTGAGAACCTTAAGCGGCAAAAGGAGCGCTCGCCCTCGCGCGCGCATGGTCC	205
Db	145	CGGGGCTTTGCAAGAACTTTGGCGGC---GAGAGGCTCTCTCTCCGCAACCGGCTGGTCC	201
Qy	206	GGGCGCTCATGCTGCACGCTGGGAAGTGTGCGATGCGAAAAACAGACGCGCGCCTCGTG	265
Db	202	GGCACTCTGTGCTCGTCTGGAAGTGTGTTGGCTGCATGTGGGAGCGCGCGTCACTCACC	261
Qy	266	CGCTTGGCGCTCCCAAGCGCGGCTTTGGAGGG-----	298
Db	262	ACGTTGTGGCTCCCTAGCGCTGGCCTTTCGGGGACCATCACAGAGCATCTTTGGCGCGC	321
Qy	299	-----ATCATCCCATCG	310
Db	322	CTCFCGCGGCTGGAGTCACTCAACTTTGCCAAACAACAGACTGTCGGCACCATCCCATCG	381
Qy	311	TCGATTGGTGAGCTTGTATCACTTCGCTATTTGGATCTCTCGGTAAATTCATTTGGTTGGG	370
Db	382	TGGATTGGTGAGCTTGCACCTTCTCTACTTGGATCTCTCACATAATTCATTTGGTTGGC	441
Qy	371	GAGGTACCAAAAAAGTTTGCAGATACGGCTCAAGAGCCTCAACACTGACGCCAGTCACTC	430
Db	442	GAGCTGCCCAA-----CCTCAAGGGGCTCCACCACAGGTCAATTTACTA	486
Qy	431	GGTATGGTTCCATTACATGCTATTGCATGTGA---GCAGTAGAAGAACGCTCGATGAA	487
Db	487	GGTATGGTTTCACTAGCATGCCAATGGATGTGAAGCTTAACAGAAGATCTTCGCCGTA	546
Qy	488	GAA CCAATACAAATATCAGGGACCAACAATAGTGTGGATCAGGGAGCAACAATTTGTT	547
Db	547	CAACCAATACAAATATCTGGGACCAACAACCTCGGTCTTATCTGGCGAAACAATCTGTG	606
Qy	548	TCGGGGAATGACACACGGTCTGATCTCGGNAATAAACAACCATGTGCTGGGAGCAACAAC	607
Db	607	TCGGGAATGACAAACATGTGTCATCTCGGGGAAACAACAATACTGTGTCTGGAGGCTTCAAC	666
Qy	608	ACTGTGTAACTGGAAGTGCAATCTGTAGTTGGTAGCAACCATGTGCTATCAGGACAC	667
Db	667	ACCGTGTAAACGGAGTGACAAATGTTTAACTGGGAGCAACCATGTCTGATCTGGGAGA	726
Qy	668	AAGCATATGTTACTGATAACAATAATGTTGTATTCGGGAAACGACAATAATGTCTGGGA	727
Db	727	AACCATATTTGTAACGACAAACAATGCTGTATCTGGGGACGCAATAATGTCTCGGG	786
Qy	728	AGCTTCATCTGTATCAGGGGACCAATAACCTGATCCGGGAGCAACAATCTGTATCC	787
Db	787	AGCTTCATAAAGTATCTGGAAGTCAACAATACGGTATCTGGGAGCAACAATACCGTATCC	846
Qy	788	GGGAGCAACCATATCGTATCTGGGAGCAACAAGTCTGTAAACAGATGGTTAAATTTCTGTA	847
Db	847	GGGAGAACCATGTCTGATCTGGGAGCAACAAGTCTGTGACAGAGGTTAATGATATGTT	906
Qy	848	GGTGCA-----GGATTGCTTCCATCTTCCCAAGTTTCAGTGTAGCT	887
Db	907	AGTGGATTGTTTCCATCTTCCTTAAGGATCTCACTGATCTTGACAAAGTTTCAGTGTAGCA	966
Qy	888	TACAAATCAATAGATGGAGACAATCACGTTATGTAACTTTCAGG-----ATATGGCATACTT	942
Db	967	CTCAATCACTTGGTGGGACAACTCGGGTTATGTAACTGATCATGGATAGCATAGTACTT	1026
Qy	943	TTCTTTTAAATAAAGCTTTCCTTTACATAAAAAAAA	980
Db	1027	TTCTCTACTTTAAATAAGAACTTTCCTATAAAAAAAA	1064

RESULT 12	
ADMD41459	ADMD41459 standard; cDNA; 959 BP.
XX AC	
XX AC	ADMD41459;
XX AC	
XX AC	03-JUN-2004 (first entry)
XX AC	
XX AC	Tall fescue antifreeze protein cDNA.
XX AC	
XX AC	Antifreeze; fescue; cold tolerance; transgenic; plant; antigitout;
XX AC	litholytic; nephrotropic; cytostatic; gene; ss.
XX AC	
XX AC	Schedonorus arundinaceus.
XX AC	
XX AC	Key
XX AC	Location/Qualifiers
XX AC	82. .813
XX AC	/tag= b
XX AC	/product= "Antifreeze protein"
XX AC	82. .141
XX AC	sig_peptide
XX AC	/tag= a
XX AC	142. .810
XX AC	mat_peptide
XX AC	/tag= c
XX AC	
XX AC	WO2004022700-A2.
XX AC	
XX AC	18-MAR-2004.
XX AC	
XX AC	09-SEP-2003; 2003WO-NZ000199.
XX AC	
XX AC	09-SEP-2002; 2002US-0409557P.
XX AC	
XX AC	(GENE-) GENESIS RES & DEV CORP LTD.
XX AC	(WRIG-) WRIGHTSON SEEDS LTD.
XX AC	
XX AC	Demmer J, Shenk MA, Hall C, Fish SA;
XX AC	
XX AC	WPI; 2004-248453/23.
XX AC	P-PSDB; ADMD41472.
XX AC	
XX AC	New antifreeze proteins and encoding polynucleotides, useful for
XX AC	modulating cold tolerance in organisms, as food additives, or for
XX AC	treating tumors or disorders associated with the presence of unw
XX AC	biocrystals (e.g. gout).
XX AC	
XX AC	Claim 1; SEQ ID NO 2; 71pp; English.
XX AC	
XX AC	The present sequence is that of cDNA encoding an antifreeze prot
XX AC	tall fescue. The cDNA was isolated from a basal stem cDNA expres
XX AC	library. The invention provides forage grass (perennial ryegrass
XX AC	fescue) antifreeze proteins and the polynucleotides encoding the
XX AC	ADMD41458-ADMD41483. The polynucleotides were isolated from tissue
XX AC	at different times of year (winter and spring) and from different
XX AC	of the plants. The polynucleotides can be used to modulate the c
XX AC	tolerance of an organism, especially plants, mammals, insects, f
XX AC	archaea and bacteria. The method involves incorporating an anti
XX AC	polynucleotide, under the control of a gene promoter sequence, i
XX AC	genome of the organism, or introducing double-stranded RNA corre
XX AC	to the polynucleotide into the cells of the organism, thereby in
XX AC	expression of an antifreeze polypeptide. The antifreeze protein
XX AC	used for the cryopreservation of a cell or tissue, as a food add
XX AC	a frozen food product, in a method for decreasing the time requi
XX AC	dehydrate a composition, to treat a disorder characterised by bi
XX AC	associated with disorders such as gout and kidney stones, to pre
XX AC	viability of a molecular biology reagent, to destroy unwanted di
XX AC	patient e.g. tumor tissue, and to protect a plant from damage d
XX AC	frost or freezing.

Best Local Similarity 70.3%; Pred. No. 6e-120;		Matches 701; Conservative 0; Mismatches 206; Indels 90; Gaps 7;	
Qy	33	CCATGCTGAATACATGGCAAGTGTTCATGCTGCTGCTCTTCTTGGGGTTCACTTTCG	92
Db	46	CCAGAACTTTAATCCATGGCGAAATGTTGGCTGCTCTTCTTGGTGTCTCTTTC	105
Qy	93	AGTGGCAGGAGCAACGTCGTGCTGTGCACACAGACGACCTCCACGCGTTGAGGGGC	152
Db	106	TGGCCATGACGGGAC-----GTCTGTGCACCTGGATGACCTCCGCGGCTGCGGGGT	159
Qy	153	TCGCTGAGAACCTTAAGCGCAAGAGCCCTCCGCTCCGCGCCGATGCTCCGCGCT	212
Db	160	TTGTCGGGAACCTCAATGCG---GGGGTGCCCTTCTCCGTGGAAATGCTCTGGCTCT	216
Qy	213	CATGCTGACGCTGGGAAGTGTGGATGCGAAACAGCAGCGCGCGCTGCTGGCGTTGC	272
Db	217	CATGCTGCGATTGGGAAGTGTGGCTGCGATGGTACAAGCGCGCGCGTCAAGCGGTTGC	276
Qy	273	GGCTCC-----CCAGCGCG	287
Db	277	GGCTCCGATTAGCTCGAGGACTGCGGTAAGCTCAAGTCGCTCAACCTTGCACAGAA	336
Qy	288	GGCTGGAGGGATCATCCCATCGTCGATTGGTGAGCTTGATCACCTTCGCTATTGGATC	347
Db	337	GATTGGTTGGCACCATCCCGTCGTGGATTGGTGAGCTTGACCACCAATTGCTACTTGGTTC	396
Qy	348	TCTCGGGTAATTCATTGGTTGGGAGGTACC-----AAAAAGTTTCGAGATAC	395
Db	397	TCTCGGATAATTCATTGGTTGGTAAGCCACCCCAATAGTTTGCACAAATAGTTTGCAGATAA	456
Qy	396	GGCTCAAGAGCCTCACCACTGACAGCAGTCACTCGGTATGGTTCCATTACATGCTAT	455
Db	457	GACTCAAGGGCCCTCGCCACCGCTGCTGTTCACTAGGTATGGCTTTCGCTAACATGCCAT	516
Qy	456	TGCATGTGA---GCAGTAGAAGAACGCTCGATCGATGAAGAACCAATAACAATATCAGGACCA	512
Db	517	TGCATGTGAAGGGACCGACGAGAACCTCGACGACAAACAATAACATACATGAGACCA	576
Qy	513	ACAATAGTTGGATCAGGAGCAACAATGTTGTTTCCGGGAATGACAAACAGCGTGTAT	572
Db	577	ACAACACTGTAGATCTGGGAACGAAATGCTGTTCTGGGAACGACAAACACTGTCTAT	636
Qy	573	CTGGGATTAACACCACTGTCTGGGAGCAACAACACTGTTGTAAGTGAAGTGACAATA	632
Db	637	GTGGGAACAACAACACTGTCTGGGAGCAACAACCAATTCGATCTGGCAGTGACAATA	696
Qy	633	CTGTAGTTGGTAGCAACCACTGCTATCAGGAGCAAGCATATTGTTACTGATAACAATA	692
Db	697	TGCTAACTGGCAGCAACCATATTGATGTGGGACCAACAATATCATATGATATACACA	756
Qy	693	ATGTTGTATCCGGGAACGAAATAATGTGTCTGGAAAGTTCCTATCTATCAGGGAGC	752
Db	757	ATGACCTATCCGCAATGATAATAATGATCTCTGGGAGCTTCCATCTATCTCCGGAGCC	816
Qy	753	ACAATACCTATCCGGGAGCAACAATATCTATCTCCGGGAGCAACCATCTGATCTGGAA	812
Db	817	ACAATACTGTATCTGGAAGTAACAACACTGTATCTGGAAGCAACCAATCTGATCTGGAA	876
Qy	813	GCAACAAAGTCGTAACAGATGGTTAATATTCTGTAGG-----TGC	852
Db	877	GCNACAACTGTGACAGAGATGAATGATTTGTACAGGGATGCTTCCATCTTTCCTAA	936
Qy	853	AGGATTGCTTCCATCTTCCCAAGTTCAGTGTAGCTTTACAAATCAATAGATGGAGACAATCA	912
Db	937	AGGAGCTCTCACCTAGTCCAAAGTTCGGTGCAGCTCACAATCACTTGGTAGGACAATCG	996
Qy	913	CGTTATGTAACCTCA--GGATATGGCACTATTTTCCTT	948
Db	997	AGTTATGTAACCTTCATGGGATATAGCATCATTTCCCT	1033

ADM41465		ADM41465 standard; cDNA; 1064 BP.	
ID	ADM41465		
AC	ADM41465;		
XX			
DT	03-JUN-2004 (first entry)		
XX			
DE	Perennial ryegrass antifreeze protein APP4 cDNA.		
XX			
KW	Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;		
KW	antigout; litholytic; nephrotropic; cyostatic; gene; ss.		
OS	Lolium perenne.		
XX			
EH	Key	Location/Qualifiers	
FT	CDS	55..897	
FT		/*tag= b	
FT		/product= "Antifreeze protein APP4"	
FT	sig_peptide	55..120	
FT		/*tag= a	
FT	mat_peptide	121..894	
FT		/*tag= c	
XX			
PN	WO2004022700-A2.		
XX			
PD	18-MAR-2004.		
XX			
PF	09-SEP-2003; 2003WO-NZ000199.		
XX			
PR	09-SEP-2002; 2002US-0409557P.		
XX			
PA	(GENE-) GENESIS RES & DEV CORP LTD.		
PA	(WRIG-) WRIGHTSON SEEDS LTD.		
XX			
PI	Demmer J, Shenk MA, Hall C, Fish SA;		
XX			
DR	WPI; 2004-248453/23.		
DR	P-PSDB; ADM41478.		
XX			
PT	New antifreeze proteins and encoding polynucleotides, useful for		
PT	modulating cold tolerance in organisms, as food additives, or for		
PT	treating tumors or disorders associated with the presence of unwanted		
PT	biocrystals (e.g. gout).		
XX			
PS	Claim 1; SEQ ID NO 8; 71pp; English.		
XX			
CC	The present sequence is that of cDNA encoding APP4, an antifreeze protein		
CC	of perennial ryegrass. The cDNA was isolated from a leaf and pseudostem		
CC	cDNA expression library. The invention provides forage grass (perennial		
CC	ryegrass and tall fescue) antifreeze proteins and the polynucleotides		
CC	encoding them ADM41458-ADM41483. The polynucleotides were isolated from		
CC	tissues taken at different times of year (winter and spring) and from		
CC	different parts of the plants. The polynucleotides can be used to		
CC	modulate the cold tolerance of an organism, especially plants, mammals,		
CC	insects, fungi, archaea and bacteria. The method involves incorporating		
CC	an antifreeze polynucleotide, under the control of a gene promoter		
CC	sequence, into the genome of the organism, or introducing double-stranded		
CC	RNA corresponding to the polynucleotide into the cells of the organism,		
CC	thereby inhibiting expression of an antifreeze polypeptide. The		
CC	antifreeze protein can be used for the cryopreservation of a cell or		
CC	tissue, as a food additive of a frozen food product, in a method for		
CC	decreasing the time required to dehydrate a composition, to treat a		
CC	disorder characterised by biocrystals associated with disorders such as		
CC	gout and kidney stones, to preserve the viability of a molecular biology		
CC	reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and		
CC	to protect a plant from damage due to frost or freezing.		
XX			
SQ	Sequence 1064 BP; 263 A; 274 C; 267 G; 260 T; 0 U; 0 Other;		

Query Match 40.2%; Score 394.4; DB 12; Length 1064;  
Best Local Similarity 67.2%; Pred. No. 1e-112;  
Matches 711; Conservative 0; Mismatches 226; Indels 121; Gaps 6;











CC different parts of the plants. The polynucleotides can be used to  
 CC modulate the cold tolerance of an organism, especially plants, mammals,  
 CC insects, fungi, archaea and bacteria. The method involves incorporating  
 CC an antifreeze polynucleotide, under the control of a gene promoter  
 CC sequence, into the genome of the organism, or introducing double-stranded  
 CC RNA corresponding to the polynucleotide into the cells of the organism,  
 CC thereby inhibiting expression of an antifreeze polypeptide. The  
 CC antifreeze protein can be used for the cryopreservation of a cell or  
 CC tissue, as a food additive of a frozen food product, in a method for  
 CC decreasing the time required to dehydrate a composition, to treat a  
 CC disorder characterised by biocrystals associated with disorders such as  
 CC gout and kidney stones, to preserve the viability of a molecular biology  
 CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
 CC to protect a plant from damage due to frost or freezing.

XX SQ Sequence 1007 BP; 252 A; 265 C; 253 G; 237 T; 0 U; 0 Other;

Query Match 50.3%; Score 492.8; DB 12; Length 1007;  
 Best Local Similarity 73.6%; Pred. No. 1.2e-143;  
 Matches 739; Conservative 0; Mismatches 197; Indels 68; Gaps 6;  
 QY 12 AATCAAGGTTCTTGTCAATCCATGCTGAATACATGCAAAAGTGTTCATGCTGCTGG 71  
 DB 37 ATTGAATCCATCTATAGCATAGCACTGCTGAATCCATGGCGAAATGCTTGATGCTGCTTC 96  
 QY 72 TCTTCTGGGGTTCATCTTTCAGGTGGCAGGA-----GCAAGTGTGCTGCTGCC 122  
 DB 97 TCTCCTTCGCTTCTTCTTGTGCGGCGCGGACGCGGACGCGGACGCCATGCC 156  
 QY 123 ACCAGCAGCACTCCACGCGTGTGAGGGGCTCGCTGAGAACCTAAGCGGCAAGAGCGCG 182  
 DB 157 ACCGGAATGACCTTCGCGGCTGCGGGGCTTCTGTGAGAACCTTGGCGGCGGCGGCGAC 216  
 QY 183 TCGCCTCCGCGCGCATGCTCGGCGCTCATGCTGAGCTGGGAGGTGTGGATGCG 242  
 DB 217 TCAGCCTCCGCGCGCGTGTGAGGCGCTCATGCTGCGATTTGGGAAGCGTGTGGCTGG 276  
 QY 243 AAACAGCAGCGCGCGCTGCTGGGCTTCCGAGCGCGCTTCCGAGCGCGCTTGGAGGATCA 302  
 DB 277 ACGGTGCCAGCGCGCGTGTGAGGCTTTCGCGTCCCGAGGAGCGCGCTCACGGGGCCAA 336  
 QY 303 TCCCATCGTGTGAGTGTGATGATCCTGCTGCTGATTTGGATCTCTCGGGTAAATTCAT 362  
 DB 337 TCCGTCATGATTTTTCAGCTTCCACCTGATGCTGCTGATCTTTGAGTATGATATGAT 396  
 QY 363 TGGTTGGGAGGTACCAAAAGTTTGCAGATAGCGCTCAAGAGCGCTCACCACTGCAGCC 422  
 DB 397 TGGTTGGGAGGTACCAAAAGTTTGCAGATAGCGCTCAAGAGCGCTCACCACTGCAGCC 446  
 QY 423 AGTCACTCGGTATGGGTTCCATTAACATGCTATTTGATGCTGTA---GCAGTAGAGNACCG 479  
 DB 447 -----AACATGCCATTGCTGATGCTGCTGATGCTGCTGATGCTGCTGAT 483  
 QY 480 TCGATGAAGAACCAATAACAATACAGGACCAACAATAGTGTGATCAGGAGGAGCAACA 539  
 DB 484 TCAGCAGCAGCCCAATACATTTCTGGGAGCAACAATAGTGTGATCAGGAGGAGCAACA 543  
 QY 540 ATGTTGTTTTCGCGGAATGACCAACCGTGTGATCTCGGGAATACCAACCATGTGTGGGA 599  
 DB 544 ATGTTTCTGCGGAATGACCAACCGTGTGATCTCGGGAATACCAACCATGTGTGTGGGA 603  
 QY 600 GCAACCAACTGTTGTAAGTGAAGTGAACATGCTAGTGTGATGCTGATGCTGCTGAT 659  
 DB 604 GCAACCAACTGTTGTAAGTGAAGTGAACATGCTAGTGTGATGCTGATGCTGCTGAT 663  
 QY 660 CAGGGAACAAGCATATTGTTACTGATAACAATATGTTGATCCGGAACGACCAATAATG 719  
 DB 664 CAGGGAACAACCATATGCTTTACAGACCAACAATAGTGTGATCAGGAGGAGCAATAATG 723  
 QY 720 TGTCTGGAAGCTTCCATCTGTTATCAGGGGAGGACCAATACCGTATCCGGGAGCAACAATA 779  
 DB 724 TATCCGGGAGCTTTCATACCGTATCCGGGGGAGGACCAATATGCTCTCCGGGAGCAACAATA 783

QY 780 CTGTATCGGAGCAACCATATCGTATCTGGAGCAACAAAGTCTGTAACAGATGCTTAAT 839  
 DB 784 CCGTATCTGGAGCAACCAACGCTTGTATCTGGAAGCAACAAAGTCTGACAGACGCTTAAT 843  
 QY 840 ATTCTGTAGGTGCGAGGATGCTTCCATTT-----CCCAAGTTGAG 880  
 DB 844 GATCTGTGAGGCGCATGATTTGTTCCACCTTAACGTGAGCTCAGTTCTTTGTCACAGTTTAC 903  
 QY 881 TGTAGCTTACATCAATAGATGAGACAAATCACGTTATGTAACCTTCA--GGATATGGCATA 939  
 DB 904 TGTACTTACAGTTCAGTGTGGTTCGCTCAATCGGTTATGTAACCTTCAATGATATACCATA 963  
 QY 940 CTTTTC-----CTTAAATAAAGCTTCCCTTTTACATAAAAAAAA 980  
 DB 964 CTTTTCCTACTATATATAAATTTCCCTTTTACATAAAAAAAA 1007

RESULT 6

ADW41463  
 ID ADW41463 standard; cDNA; 1230 BP.

XX AC ADW41463;

XX DT 03-JUN-2004 (first entry)

XX DE Tall fescue antifreeze protein cDNA.

XX KW Antifreeze; fescue; cold tolerance; transgenic; plant; antigout;

XX KW litholytic; nephrotropic; cytotstatic; gene; ss.

XX OS Schedonorus arundinaceus.

XX FH Key Location/Qualifiers

XX FT CDS 76..909

XX FT /tag= b

XX FT /product= "Antifreeze protein"

XX FT sig\_peptide 76..141

XX FT /tag= a

XX FT mat\_peptide 142..906

XX FT /tag= c

XX WO2004022700-A2.

XX PN 18-MAR-2004.

XX PD 09-SEP-2003; 2003WO-NZ0000199.

XX PR 09-SEP-2002; 2002US-0409557P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PA (WRIG-) WRIGHTSON SEEDS LTD.

XX PI Demmer J, Sherk MA, Hall C, Fish SA;

XX XX WPI: 2004-248453/23.

XX DR P-PSDB; ADW41476.

XX PT New antifreeze proteins and encoding polynucleotides, useful for  
 PT modulating cold tolerance in organisms, as food additives, or for  
 PT treating tumors or disorders associated with the presence of unwanted  
 PT biocrystals (e.g. gout).

XX PS Claim 1; SEQ ID NO 6; 71pp; English.

XX CC The present sequence is that of cDNA encoding an antifreeze protein of  
 CC tall fescue. The invention provides forage grass (perennial ryegrass and  
 CC tall fescue) antifreeze proteins and the polynucleotides encoding them  
 CC ADW41458-ADW41483. The polynucleotides were isolated from tissues taken  
 CC at different times of year (winter and spring) and from different parts  
 CC of the plants. The polynucleotides can be used to modulate the cold  
 CC tolerance of an organism, especially plants, mammals, insects, fungi,  
 CC archaea and bacteria. The method involves incorporating an antifreeze  
 CC polynucleotide, under the control of a gene promoter sequence, into the

rygrass and tall fescue) antifreeze proteins and the polynucleotides encoding them ADM4148-ADM41483. The polynucleotides were isolated from tissues taken at different times of year (winter and spring) and from different parts of the plants. The polynucleotides can be used to modulate the cold tolerance of an organism, especially plants, mammals, insects, fungi, archaea and bacteria. The method involves incorporating an antifreeze polynucleotide, under the control of a gene promoter sequence, into the genome of the organism, or introducing double-stranded RNA corresponding to the polynucleotide into the cells of the organism, thereby inhibiting expression of an antifreeze polypeptide. The antifreeze protein can be used for the cryopreservation of a cell or tissue, as a food additive of a frozen food product, in a method for decreasing the time required to dehydrate a composition, to treat a disorder characterised by biocrystals associated with disorders such as gout and kidney stones, to preserve the viability of a molecular biology reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and to protect a plant from damage due to frost or freezing.

Sequence 1008 BP; 256 A; 265 C; 251 G; 236 T; 0 U; 0 Other

SQ Sequence 1008 BP; 256 A; 265 C; 251 G; 236 T; 0 U; 0 Other;

Query Match	51.4%	Score 503.6;	DB 12;	Length 1008;
Best Local Similarity	74.3%	Pred. No. 4.7e-147;		

12	Qy	AATCAAGGTTTCTGTTCAATCCATGCGCTGAATACATGCCAAGTGTGTGATGCTGCTGG 71
44	Db	ATTGAATCCATCTATAGCATAGCATGCTGTAATCCATGGGAAATGCTTGATGCTGCTGC 103
72	Qy	TCCTCTTGGGGTTCACTCTGCAGGTGGCAGGA---GCAACGTGCTGCTGCTGCCACGAG 128
104	Db	TTCTCTTTCGGTTCCTCTTTGTCGGTGGCCGCGACGGCGACGGCGAGCCCATGCCACGGG 163
129	Qy	AGACCTCCACACGCTTGAGGGGCCCTCGCTGAGAACTTAAGCGGCAAGAGAGCCGTCCGCC 188
164	Db	ATGACCTTCGCGCGCTGCGGGGCTTCGCTGAGAACCTTGGGCGCGCGCGCAATCAGCC 223
189	Qy	TCCGCGCCGCAATGGTCCGCGCCTCATGCTGACGTGGGAAGGTGGGATGCGAAACAG 248
224	Db	TCCGCGCCGCTGGTTCAGGCGCCTCATGCTCGAATTGGGAAGGCGTTGGCTGCGACGGTG 283
249	Qy	CAAGCGGCGCGCTGCTGGGGTTTGCGGCTCCCAAGCGGGGCTTGGAGGATCATCCCAT 308
284	Db	CCAGCGGCGCGTCTCAGGCTTTGTGGCTCCCGAGGAGCGGCCTCACGGGGCCAATCCCGT 343
309	Qy	CGTCGAATTGGTCAGCTTGATCACTTCGCTATTTTGGATCTCTCGGGTAATTCAATCGGTG 368
344	Db	CATGGAATTGTGTCAGCTTCCACACCTACGCTACTTGGATCTTTTCAAGTAATGCAATGGTTG 403
369	Qy	GGGAGGTACCAAAAAGTTTGCAGATACGGCTCAAGAGCCTCAACCACTGACGACGAGTCAC 428
404	Db	GCAGGTACCCAGAATCTGCAGGTACAGCTCAAAGGCATCACC-----447
429	Qy	TCGGTATGGGTTCCAAATTAACTGCTAATTGCAATGTGA---GCAGTGAAGAAACGCTCGATG 485
448	Db	-----AACATGCCAATTGCATGTGATCGCTTAAAGAAAGATCACTCGACG 490
486	Qy	AAGAACCAATAACAATATCAGGGCCACCAATAGTGTGGATCAGGGAGCACAATGTGTG 545
491	Db	AGAGCCCAATAACAATTTCTGGAGACAAATACTGTCCAGTCCGGAGCAAAAATGTTC 550
546	Qy	TTTCCGGGAATGACAAACCGTTCGTATCTGGGAATAACAACCATGTGTCTGGGAGCAACA 605
551	Db	TTGCTGGGAATGACAAACCGTCATATCTGGGACACAATAGTGTGTCTGGAGCAACA 610
606	Qy	ACACTGTTGTAACTGGGAAGTGACAATACTGTAGTGTGGTATGGAACCAATGTGTAATCAGGA 665
611	Db	ACACTGTGCTAAGTGGGAATGACAATAACCGTAAACCGGCACCAACCATGTGTAATCAGGA 670
666	Qy	CAAGCATATTGTTACTGATACCAATAATGTTGTATCCGGGAACGACAATAATGTGTCTG 725
671	Db	CAAACCATATCGTTTACAGACAAACAATAACGTAATCCGGGAACGATAATAATGTATCCG 730
726	Qy	GAAGCTTCCATCTGTATCAGGGGAGACAATAACCGTATCCGGGAGCAACAATACTGTAT 785



CC The present sequence is that of cDNA encoding an antifreeze protein of  
CC tall fescue. The cDNA was isolated from a leaf blade cDNA expression  
CC library. The invention provides forage grass (perennial ryegrass and tall  
CC fescue) antifreeze proteins and the polynucleotides encoding them  
CC ADMA1458-ADMA1483. The polynucleotides were isolated from tissues taken  
CC at different times of year (winter and spring) and from different parts  
CC of the plants. The polynucleotides can be used to modulate the cold  
CC tolerance of an organism, especially plants, mammals, insects, fungi,  
CC archaea and bacteria. The method involves incorporating an antifreeze  
CC polynucleotide, under the control of a gene promoter sequence, into the  
CC genome of the organism, or introducing double-stranded RNA corresponding  
CC to the polynucleotide into the cells of the organism, thereby inhibiting  
CC expression of an antifreeze polypeptide. The antifreeze protein can be  
CC used for the cryopreservation of a cell or tissue, as a food additive of  
CC a frozen food product, in a method for decreasing the time required to  
CC dehydrate a composition, to treat a disorder characterised by biocrystals  
CC associated with disorders such as gout and kidney stones, to preserve the  
CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
CC patient e.g. tumour tissue, and to protect a plant from damage due to  
CC frost or freezing.

XX SQ Sequence 1006 BP; 248 A; 261 C; 260 G; 237 T; 0 U; 0 Other;

Query Match 55.3%; Score 541.6; DB 12; Length 1006;  
Best Local Similarity 76.5%; Pred. No. 5.3e-159;  
Matches 756; Conservative 0; Mismatches 194; Indels 38; Gaps 6;

QY 22 TCTTGTTCATCATCGCTGAATACATGCGCAAGTGTTCATGCTGCTGCTTCTTGGG 81  
DB 28 TCTATAGCATAGACATGCTGTAATCATGCGCAAAATGCTTGATGCTGCTCTCTCGC 87  
QY 82 GTTCATCTTGCAGGTGGCAGGA---GCAACGTCGTGCTGTCGCCACACGACCTCCA 138  
DB 88 GTTCTCTTGTGCGCGGCTGGCAGCGCGACGCGCATGCCACCGGATGACTCGG 147  
QY 139 CGCGTTGAGGGGCTCGCTGAGAACCTAAGCGCAAGAGCGCTCCGCTCCGCGCGCG 198  
DB 148 TCGCTGCGGGGCTTCGCTGAGAACCTGCGCGCGCGCGCACTCAGCTTCGCGCGCG 207  
QY 199 ATGGTCCGCGGCTCATGCTGAGCTGGGAAGTGTGGATGCGAAACGAAAGCGCGG 258  
DB 208 GTGGTCAGCGGCTCATGCTGCGATTGGGAAGCGTGTGGCTGCGACGCTGCCAGCGCG 267  
QY 259 CGTCTGGGCTTGGCGCTCCCAAGCGGCTTGGAGGATCATCCCATCTCTCATTTGG 318  
DB 268 TGTACGGCTTTGTGGCTCCCGAGGAGCGGCTCAACGCGGCGCAATCCCATCTGTGATTG 327  
QY 319 TGAGCTTTGATCACTTCGCTATTGGATCTCTCGGTAATTCATTGTTGGGAGGTACC 378  
DB 328 TGAGCTTACCACTACGCTATTGGATCTTTGAGTAATGATGTTGGCGAGGTACC 387  
QY 379 AAAAAATTTGCGATACGCTCAAGAGCTCAACCACTGACAGCGAGTCACTCGGTATGGG 438  
DB 388 CAAGATCTGAGGTACAGCTCAAGGCGCTCACCGCTGCGGTC-----GTTGGG 438  
QY 439 TTCCATTACATGCTATTGATGTGA---GCAGTAGAAGAACGCTCGATGAGAACCAAA 495  
DB 439 TTTCAACAACATGCCATTGCTGATGTGATGCGTAACAGAGATCACTCGACGAGCGCCAA 498  
QY 496 TACAAATATCAGGACCAACAATAGTCTTGGATCAGGAGCAACAATGTTGTTCCGGGAA 555  
DB 499 TACAATATCTGGGAGCAACAATACTGTCTAGATCCGGGAGCAAAAATGTTGTTCTGGGAA 558  
QY 556 TGACAAACAGGCTCGTATCTGGGAATAACAACCACTGTTCTGGGAGCAACAACACTGTTGT 615  
DB 559 TGACAAACAGGCTCATATCTGGGAGCAACAATACTGTTCTGGGAGCAACAACACTGTTGT 618  
QY 616 AACTGGAAGTGACAATCTGTTAGTGTGATGCAACCATGCTGATCAGGAGCAACAACATAT 675  
DB 619 AAGTGGGAGTGACAATACCGTAACTGCGACCAACCATGCTGATCAGGAGCAACAACATAT 678  
QY 676 TGTACTGTATACAATAATGTTCTATCCGGGAGCAACAATAATGCTCTGGAGCTTCCA 735

DB 679 CGTTACAGACAAACAATAACGTTATCCGGGAACGATAAATATGTTATCCGGAGCTTTCA 738  
QY 736 TACTGTATCAGGGAGGACAAATACGTTATCCGGAGCAACAATACTGTATCCGGAGCAAA 795  
DB 739 TACCGTATCCCGGGGCAAAATACCGTCTCTGGGAGCAACAATACTGTATCTGGAGCA 798  
QY 796 CCATATCTGATCTCTGGGAGCAACAAGTCTGAACAGATGGTTAATATCTGTAGGTGCGAG 855  
DB 799 CCAATGTTGATCTGGAAGCAACAAGTCTGACAGACGCTTAATGATCTGTGAGGCGATG 858  
QY 856 ATTGCTTCCATCTT-----CCCAAGTTCAGTGTAGCTTACAATCAA 896  
DB 859 ATTTGTTCCACCTTAACTGAGCTCAGTCTTGTGTCGAAGTTCACGTACCTCACAGTCA 918  
QY 897 TAGATGGAGACAATCACGTTATGTAACCTTCA-GGATATGGCATACTTTTC---CTTTAAA 952  
DB 919 TTGGTGGCTTCAATCGGCTTATGTAACCTTCATGATATACCATATCTTTTCTACTATA 978  
QY 953 TAAAGCTTCCCTTTTACATAAAAAA 980  
DB 979 TAAATTTCCCTTTAAAAA 1006

RESULT 4  
ADMA1467

ID ADMA1467 standard; cDNA; 1008 BP.

XX AC ADMA1467;

XX XX 03-JUN-2004 (first entry)

XX DE Perennial ryegrass antifreeze protein APP5 cDNA.

XX XX Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
KW antitox; litholytic; nephrotropic; cytostatic; gene; ss.

XX OS Lolium perenne.

XX FH Key Location/Qualifiers  
CDS 79..843

FT /\*tag= b /product= "Antifreeze protein APP5"

FT sig\_peptide 79..147

FT /\*tag= a

FT mat\_peptide 148..840

FT /\*tag= c

XX WO2004022700-A2.

XX PN 18-MAR-2004.

XX PF 09-SEP-2003; 2003WO-NZ000199.

XX PR 09-SEP-2002; 2002US-0409557P.

XX XX (GENE-) GENESIS RES & DEV CORP LTD.

XX PA (WRIG-) WRIGHTSON SEEDS LTD.

XX XX Demmer J, Shenk MA, Hall C, Fish SA;

XX XX WPI: 2004-248453/23.

XX DR P-PSDB; ADMA1480.

XX XX New antifreeze proteins and encoding polynucleotides, useful for  
modulating cold tolerance in organisms, as food additives, or for  
treating tumors or disorders associated with the presence of unwanted  
biocrystals (e.g. gout).

XX PS Claim 1; SEQ ID NO 10; 71pp; English.

XX CC The present sequence is that of cDNA encoding APP5, an antifreeze protein  
of perennial ryegrass. The cDNA was isolated from a leaf blade cDNA  
expression library. The invention provides forage grass (perennial

Claim 1; SEQ ID NO 4; 71pp; English.  
 The present sequence is that of cDNA encoding an antifreeze protein of tall fescue. The cDNA was isolated from a leaf blade cDNA expression library. The invention provides forage grass (perennial ryegrass and tall fescue) antifreeze proteins and the polynucleotides encoding them (ADW41458-ADW41483). The polynucleotides were isolated from tissues taken at different times of year (winter and spring) and from different parts of the plants. The polynucleotides can be used to modulate the cold tolerance of an organism, especially plants, mammals, insects, fungi, archaea and bacteria. The method involves incorporating an antifreeze polynucleotide, under the control of a gene promoter sequence, into the genome of the organism, or introducing double-stranded RNA corresponding to the polynucleotide into the cells of the organism, thereby inhibiting expression of an antifreeze polypeptide. The antifreeze protein can be used for the cryopreservation of a cell or tissue, as a food additive or a frozen food product, in a method for decreasing the time required to dehydrate a composition, to treat a disorder characterised by biocrystals associated with disorders such as gout and kidney stones, to preserve the viability of a molecular biology reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and to protect a plant from damage due to frost or freezing.

Seq	Sequence	996 BP;	268 A;	235 C;	250 G;	243 T;	0 U;	0 Other;
Query Match	92.3%;	Score	904.4;	DB	12;	Length	996;	
Best Local Similarity	96.2%;	Pred.	No. 5.2e-273;					
Matches	949;	Conservative	0;	Mismatches	31;	Indels	6;	Gaps
Qy	1	GCTTGCATTC	CAATCAAGGTTTCTTGTTCAATCCATGCTGAATACATGGCAAGTGTTG	60				
Db	7	GCTTGCATTC	CAATCAAGGTTTCTTGTTCAATCCATGCTGAATACATGGCAAGTGTTG	66				
Qy	61	CATGCTGCTG	CTCTCTTGGGGTTCA---TCTTCAGGTGGCAGGAGCAACCTCGTGTGTC	117				
Db	67	CATGCTGCTG	CTCTCTTGGGGTTCAATCTTTCAGGTGGCAGGAGCAACCTCGTGTGTC	126				
Qy	118	GTGCCACCA	CAGCAGCACTCCACGGCTTCAGGGGCTCGCTGAGAACTTAAGCGGCAAGG	177				
Db	127	GTGCCACCA	CAGCAGCACTCCCGGCATTCAGGGGCTTCGCGNGAACCTTAAGCGGCAAGG	186				
Qy	178	AGCCGTCGC	CTCCGCGCCGATGATGTCGGCGCCCTCATGCTGCAGCTGGGAAGTGTGGG	237				
Db	187	AGCCGTCGC	CTCCGCGCCGATGATGTCGGCGCCCTCATGCTGCAGCTGGGAAGTGTGGG	246				
Qy	238	ATGCGAAAC	ACAGACGGCCGCGTCTGGTGGTTCACCTTCGCTATTTGGATCTCTCGGGTAA	357				
Db	247	ATGCGAAAC	ACAGACGGCCGCGTCTGGTGGTTCACCTTCGCTATTTGGATCTCTCGGGTAA	366				
Qy	298	GATCATCC	ATCGTTCGATTGAGCTTGATCACCTTCGCTATTTGGATCTCTCGGGTAA	357				
Db	307	GACCATCC	ATCGTTCGATTGAGCTTGATCACCTTCGCTATTTGGATCTCTCGGGTAA	366				
Qy	358	TTCAATGGT	TGGGGAGGTACCAAAAAGTTTGACATACGGCTCAAGAGCCTCACCACTGA	417				
Db	367	TTCAATGGT	TGGGAAGGTACCAAAAAGTTTGACATACGGCTCAAGAGCCTCACCACTGA	426				
Qy	418	CAGCCAGT	CACTCGGTATGGGTTCATTAACATGCTATTGCATGTGAGCAGT---AGAG	474				
Db	427	TGGCCAGT	CACTCGGTATGGGTTCATTAACATGCTATTGCATGTGAGCAGTAAACAGAG	486				
Qy	475	AACGCTCG	ATGAAGAACCAATACAAATACAGGGACCAACAATAGTGTGGATCAGGGAG	534				
Db	487	AACGCTCG	ATGAAGAACCAATACAAATACAGGGACCAACAATAGTGTGGATCAGGGAG	546				
Qy	535	CAA CAATG	TGTTTTCGGGGAATGACAAACCGGTGCTATCTCGGGAATAACCAACCATGTGTC	594				
Db	547	CAA CAATG	TGTTTTCGGGGAATGACAAACCGGTGCTATCTCGGGAATAACCAACCATGTGTC	606				
Qy	595	TGGGAGCA	ACAACACTGTTTAACTGGGAAGTGACAATCTGTAGTGTGGTAGCAACCATGTG	654				
Db	607	TGGGAGCA	ACAACACTGTTTAACTGGGAAGTGACAACACTTTAGTGTGGTAGCAACCATGTG	666				



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2005, 10:18:05 ; Search time 572 Seconds  
(without alignments)  
10142.214 Million cell updates/sec

Title: US-10-657-852B-3  
Perfect score: 980  
Sequence: 1 gcttcattcccaatcaagtg.....ccctttacataaaaaaaaaa 980

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	980	100.0	980	12 ADM41460	Adm41460 Perennial
2	904.4	92.3	996	12 ADM41461	Adm41461 Tall fesc
3	541.6	55.3	1006	12 ADM41468	Adm41468 Tall fesc
4	503.6	51.4	1008	12 ADM41467	Adm41467 Perennial
5	492.8	50.3	1007	12 ADM41469	Adm41469 Perennial
6	451.2	46.0	1230	12 ADM41463	Adm41463 Tall fesc
7	443.6	45.3	1212	12 ADM41464	Adm41464 Perennial
8	421.8	43.0	1246	12 ADM41466	Adm41466 Tall fesc
9	419	42.8	1083	12 ADM41470	Adm41470 Perennial
10	417.4	42.6	1084	12 ADM41462	Adm41462 Perennial
11	394.4	40.2	1064	12 ADM41465	Adm41465 Perennial
12	270.4	27.6	959	12 ADM41459	Adm41459 Tall fesc
13	264.2	27.0	357	2 AAX99717	Aax99717 Grass ant
14	262.2	26.8	841	12 ADM41458	Adm41458 Perennial
15	79.4	8.1	4536	12 ADJ40496	ADJ40496 Plant CDN
16	54	5.5	2000	8 ADA71938	Ada71938 Rice gene
17	51.8	5.3	1138	1 AAN91630	Aan91630 3' sequen
18	50.6	5.2	1912	12 ADJ39568	ADJ39568 Plant CDN
19	48.2	4.9	3000	12 ADI32615	ADI32615 Rice tran
20	44.8	4.6	3453	3 AAA97909	Aaa97909 Eucalyptu

21	43.8	4.5	235070	11	ACN45174	Acn45174 Human gen
C	43	4.4	7722	3	AAA70168	Aaa70168 Plasmodiu
22	41.6	4.2	2274	4	ABL17081	Abli17081 Drosophil
23	41.6	4.2	4274	4	ABL17080	Abli17080 Drosophil
C	41.6	4.2	966	6	ABQ17272	Abqi17272 Oligonuct
24	41.4	4.2	966	6	ABQ17273	Abqi17273 Oligonuct
25	41.4	4.2	408	3	AAA79411	Aaa79411 Eucalyptu
26	41.4	4.2	408	3	AAA79411	Aaa79411 Eucalyptu
27	41	4.2	702	5	AA667541	Aaa667541 DNA encod
C	40.8	4.1	396	3	AAA79420	Aaa79420 Eucalyptu
28	40.4	4.1	12614	4	AAI99322	Aai99322 Human exc
29	40.4	4.1	12614	4	AAI99322	Aai99322 Human exc
30	40.4	4.1	3228	5	AA81183	Aas81183 DNA encod
C	40.4	4.1	804	12	ADM80069	Adm80069 Spiramyci
31	40	4.1	804	12	ADM80069	Adm80069 S ambofac
32	40	4.1	30943	12	ADM80034	Adm80034 Spiramyci
33	40	4.1	30943	12	ADM80034	Adm80034 S ambofac
C	39.6	4.0	2000	8	ADA71938	Ada71938 Rice gene
C	39.6	4.0	2298	4	ABL18481	Abli18481 Drosophil
34	39.6	4.0	2479	11	ADM03281	Adm03281 Human CDN
35	39.6	4.0	4298	4	ABL18480	Abli18480 Drosophil
36	39.2	4.0	1443	10	ABZ23668	Abz23668 H. pylori
37	39.2	4.0	2274	4	ABL17081	Abli17081 Drosophil
C	39.2	4.0	4274	4	ABL17080	Abli17080 Drosophil
C	38.8	4.0	1028	3	AAA50251	Aaa50251 Maize hea

ALIGNMENTS

RESULT 1  
ADM41460  
ID ADM41460 standard; cDNA; 980 BP.  
XX  
AC ADM41460;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Perennial ryegrass antifreeze protein AFP2 cDNA.  
XX  
KW Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
KW antigout; litholytic; nephrotropic; cytotstatic; gene; ss.  
XX  
OS Lolium perenne.  
XX  
FH Key Location/Qualifiers  
FT CDS 35..838  
FT /\*tag= b  
FT /product= "Antifreeze protein AFP2"  
FT sig\_peptide 35..103  
FT /\*tag= a  
FT mat\_peptide 104..835  
FT /\*tag= c  
XX  
XX WO2004022700-A2.  
XX  
XX PN 18-MAR-2004.  
XX  
XX PD 09-SEP-2003; 2003WO-NZ000199.  
XX  
XX PF 09-SEP-2002; 2002US-0409557P.  
XX  
XX PR (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX PA (WRIG-) WRIGHTSON SEEDS LTD.  
XX  
XX PI Demmer J, Shenk MA, Hall C, Fish SA;  
XX  
XX DR WPI; 2004-248453/23.  
XX  
XX DR P-PSDB; ADM41473.  
XX  
XX PT New antifreeze proteins and encoding polynucleotides, useful for  
XX modulating cold tolerance in organisms, as food additives, or for  
XX treating tumors or disorders associated with the presence of unwanted

**This Page Blank (uspto)**

```
Db 172320 AGCATGTATCAGCACTGGGGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTTT 172261
Qy 760 CGTATCCGGGAGCAACAATACTGTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAA 819
Db 172260 TGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAACATGTATCAGCAT 172201
Qy 820 AGTCGTAACA 829
Db 172200 TGGTATAACA 172191
```

Search completed: October 8, 2005, 19:05:38  
Job time : 4305 secs

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GVQS  
Center clone name: CH230-462N7  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 165932 bases at least Q40  
Consensus quality: 167252 bases at least Q30  
Consensus quality: 167961 bases at least Q20  
Estimated insert size: 169210; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1 172927: contig of 172927 bp in length.

----- Location/Qualifiers  
1. 172927  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-462N7"  
1. 1254  
/note="wgs end extension  
clone\_end:Sp6"  
1305\_-2525  
/note="wgs end extension  
clone\_end:Sp6"  
3190\_-4089  
/note="clone boundary  
clone\_end:Sp6"  
site:  
end sequence:BZ144778"  
complement(167491..168569)  
/note="clone boundary  
clone\_end:T7"  
site:  
end sequence:BZ144777"  
169255\_-172927  
/note="wgs end extension  
clone\_end:T7"

FEATURES  
source  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
ORIGIN  
Query Match 6.1%; Score 59.6; DB 2; Length 172927;  
Best Local Similarity 52.4%; Pred. No. 0.00021;  
Matches 131; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
QY 580 TAACAACCATGTGTGGGAGCAACAACACTGTTTGTAACTGGAAGTGAACATCTGTAGT 639  
DB 172440 TATCAGCACTGTGGTAAACATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAGC 172381  
QY 640 TGGTAGCAACCATGTCTGTATCAGGACAAACCATATGTTTACTGATACATAATGTTGT 699  
DB 172380 ATGTATCAGCACTGTGTGTAGCATGTATCAGCACTGTGTGTATATGTATCAGCACTGTTGT 172321  
QY 700 ATCCGGGAACCAATAATGTGTCTGGAAGCTTCCATACTGTATCAGGGGAGCAATAC 759

Rattus.  
1 (bases 1 to 172927)  
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,  
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,  
Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,  
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,  
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,R., Garza,M.,  
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,  
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogues,M.,  
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,  
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
Kovis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
Lorensuhewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,  
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
Morgan,M., Morris,K., Morris,S., Munitasa,M., Murphy,M., Nair,L.,  
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
Nwackelme,O., Okwou,G., Olarnpunsagoon,A., Pal,S., Parks,K.,  
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,  
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,  
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,  
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,  
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,  
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,  
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
Weinstock,G. and Gibbs,R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 172927)  
Worley,K.C.  
Direct Submission  
Submitted (01-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 172927)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaesana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Unpublished  
2 (bases 1 to 261587)  
Worley, K.C.

TITLE  
JOURNAL

REFERENCE  
AUTHORS  
JOURNAL

Submitted (04-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 261587)  
Rat Genome Sequencing Consortium.

TITLE  
JOURNAL

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GHUJ  
Center clone name: CH230-104014  
----- Summary Statistics

Assembly program: Atlas 3.0;  
Consensus quality: 245312 bases at least Q40  
Consensus quality: 249138 bases at least Q30  
Consensus quality: 252151 bases at least Q20  
Estimated insert size: 257796; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html))  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 250010: contig of 250010 bp in length  
\* 250011 250110: gap of unknown length  
\* 250111 253467: contig of 3357 bp in length  
\* 253468 253567: gap of unknown length

\* 253568 255148: contig of 1581 bp in length  
\* 255149 255248: gap of unknown length  
\* 255249 256807: contig of 1559 bp in length  
\* 256808 256907: gap of unknown length  
\* 256908 258585: contig of 1678 bp in length  
\* 258586 258685: gap of unknown length  
\* 258686 260166: contig of 1481 bp in length  
\* 260167 260266: gap of unknown length  
\* 260267 261587: contig of 1321 bp in length.

#### FEATURES

source  
1..261587  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-104014"

misc\_feature

1..2511  
/note="wgs\_end\_extension  
clone end:77"  
2562..3793  
/note="wgs\_end\_extension  
clone end:77"

misc\_feature

6414..7318  
/note="clone\_boundary  
clone end:77  
site:EcORI  
end\_sequence: BH325436"

misc\_feature

complement(246622..247488)  
/note="clone\_boundary  
clone end:Sp6  
site:EcORI  
end\_sequence: BH325437"

misc\_feature

#### ORIGIN

Query Match 6.2%; Score 60.8; DB 2; Length 261587;  
Best Local Similarity 51.9%; Pred. No. 0.00011;  
Matches 165; Conservative 0; Mismatches 147; Indels 6; Gaps 1;  
  
QY 512 AACATAGTGTGGATCAGGAGCAACATGTTTCGGGAATGACACACGGTCGTA 571  
DB 83753 ATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTTTTGTAGCA 83694  
  
QY 572 TCTGGGAATAACAACCATGTGCTCGGAGCAACAACACATGTTGTAACATGGAAGTGACAAT 631  
DB 83693 T-----GTATCAGCACTGTGTTAATGATATCAGCACTGTTGTAGCATGTATCAGCACT 83640  
  
QY 632 ACTGTAGTGTGGTAGCAACCATGTCGTATCAGGAGCAAAAGCATATTGTTACTGATAACAAAT 691  
DB 83639 GTTGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTGTTAATATGATATCAGC 83580  
  
QY 692 AATGTTGTATCCGGGAACGACAAATATGTTCTGGAAGTTCCATCACTGTATCAGGGGAG 751  
DB 83579 ACTGTTGTAGCATGTATCAGCACTGTGTTAATGATATCAGCACTGTTGTAGCATGTATC 83520  
  
QY 752 CACATACCGTATCCGGAGCAACATATCTGTATCCGGAGCAACCATATCGTACTCTGG 811  
DB 83519 AGCACTTTTGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAAATGT 83460  
  
QY 812 AGCAACAAGTCGTAACA 829  
DB 83459 ATCAGCATGTTGTAACA 83442

#### RESULT 15

AC119701/c  
LOCUS AC119701 172927 bp DNA linear HTG 20-NOV-2002  
DEFINITION Rattus norvegicus clone CH230-462N7, WORKING DRAFT SEQUENCE.  
ACCESSION AC119701  
VERSION AC119701.7 GI:25137804  
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;









```

10 - 19 : 48
20 - 29 : 165
30 - 39 : 825
40 - 49 : 3919
50 - 59 : 6201
60 - 69 : 8574
70 - 79 : 22106
80 - 89 : 54448
90 - 99 : 79950
-----
Percentage of bases with a quality value >= 40 : 99 %.
FEATURES
source
  Location/Qualifiers
    1..176237
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="14"
      /clone="R-731F5"
      /clone_lib="RPCI-11"
      /note="matching EMBL:G14654"
      14470..14567
      /note="matching EMBL:G14654"
      RHdb:RH7974
      dbSTS:STS21065
      Identified using the e-PCR software (G. Schuler) "
      57606..57752
      /note="matching EMBL:L00022"
      RHdb:RH68981
      dbSTS:STS48887
      Identified using the e-PCR software (G. Schuler) "
      60040..60260
      /note="matching EMBL:M55420"
      RHdb:RH80278
      dbSTS:STS56700
      Identified using the e-PCR software (G. Schuler) "
      69316..69421
      /note="matching EMBL:H26187"
      RHdb:RH44850
      dbSTS:STS37918
      Identified using the e-PCR software (G. Schuler) "
      145938..146060
      /note="matching EMBL:253040"
      RHdb:RH9751
      dbSTS:STS5550
      Identified using the e-PCR software (G. Schuler) "
      158529..158653
      /note="matching EMBL:AA160692"
      RHdb:RH48485
      dbSTS:STS41544
      Identified using the e-PCR software (G. Schuler) "
ORIGIN
Query Match 7.2%; Score 70.2; DB 9; Length 176237;
Best Local Similarity 56.5%; Pred. No. 2.8e-07;
Matches 130; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 600 GCAACAACTGTGTAACTGGAAGTGACAATACTAGTTGTGTAGCAACCATGTCGTAT 659
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7784 GAAGCATCCAGATGGAAACAGGAAATGTCAGCATGGAACAGGAGCATCCAGAAATGTA 7725
QY 660 CAGGGCAAGCATATTTGTTACTGATAACAATAATGTTGTTATCCGGAGACGACAAATG 719
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7724 CAGGAAGCATCCAGATGGAACAGGAAATGTCAGCATGGAACAGGAGCATCAAGAAATG 7665
QY 720 TGTCTGGAAGCTTCCATCTGTATCAGGGAGACACAATACCGTATCCGGAGCAACAATA 779
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7664 TAACAGAGATCCAGATGGAGTGGACAGGAAATGTCAGCATGGAACAGGAGCATCCAGA 7605
QY 780 CTGTATCCGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTAACA 829
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7604 ATGTAACAGAGATCCAGCATGGAACAGGAGCATCCAGATGGAACA 7555
-----
CNS01DT2 169802 bp DNA linear PRI 19-NOV-2001
LOCUS Human chromosome 14 DNA sequence BAC R-417P24 of library RPCI-11
DEFINITION Human chromosome 14 of Homo sapiens (human), complete sequence.
ACCESSION AL122127
VERSION AL122127.6 GI:17026193
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 169802)
AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissensbach,J.
Sequencing of the human chromosome 14
Unpublished
REFERENCE 2 (bases 1 to 169802)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT On Nov 20, 2001 this sequence version replaced gi:14715169.
-----
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
-----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.56x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 :
1 - 9 :
10 - 19 :
20 - 29 : 11
30 - 39 : 159
40 - 49 : 4325
50 - 59 : 11964
60 - 69 : 13264
70 - 79 : 28113
80 - 89 : 55713
90 - 99 : 56253
-----
Percentage of bases with a quality value >= 40 : 99 %.
FEATURES
Location/Qualifiers
  1..169802
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="14"
    /clone="R-417P24"
    /clone_lib="RPCI-11"
    11427..11641
    /note="matching EMBL:G33053"
    RHdb:RH67749
    dbSTS:STS47676
    Identified using the e-PCR software (G. Schuler) "
    11439..11536
    /note="matching EMBL:G14654"
    RHdb:RH7974
    dbSTS:STS21065
    Identified using the e-PCR software (G. Schuler) "
    39191..39405
    /note="matching EMBL:G33053"
    RHdb:RH67749
  STS
  STS
  STS

```







[illegible]





SOURCE  
ORGANISM

Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
AUTHORS

1 Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (G33) genomic DNA, chromosome 7, PAC  
clone: P0585H11

## JOURNAL

Published Only in Database (2001)  
2 (bases 1 to 149371)

REFERENCE  
AUTHORS

Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission

## JOURNAL

Submitted (14-NOV-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: tsasaki@nias.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>,  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

## COMMENT

On Jul 22, 2004 this sequence version replaced gi:34395173.

Genes were predicted from the integrated results of the following:

GENSCAN (<http://ccr-081.mit.edu/GENSCAN.html>), FGENESH  
(<http://www.softberry.com/>), GeneMark.hmm  
(<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM  
(<http://www.tigr.org/tdb/glimmer/glmr form.html>), RiceHMM  
(<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor  
(<http://bioinformatics.fastate.edu/cgi-bin/sp.cgi>), sim4  
(<http://globin.cse.psu.edu/html/docs/sim4.html>), gap2  
(<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The  
genomic sequence was searched against NCBI Nonredundant Protein  
database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA  
sequence database at RGP or DDBJ. Protein homologs of the coding  
regions were searched against NCBI Nonredundant Protein database  
with BLASTP. ESTs represent the identified cDNA sequences using  
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.  
Full-length cDNAs represent the identified cDNA sequences using  
BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level  
such as same name, 'putative-', and '-like protein'. A gene without  
significant homology to any protein but with full-length cDNA or  
EST homology (covering almost the entire length of partial  
sequence) is classified as an 'unknown' protein. A gene predicted  
by two or more gene prediction programs is classified as a  
'hypothetical' protein according to IKGSP standard. A gene  
predicted by a single gene prediction program is also classified as  
a probable 'hypothetical' protein and is included as a  
miscellaneous feature of the sequence.

The orientation of the sequence is from T7 to SP6 of the PAC clone.  
This sequence of P0585H11 clone has an overlap with B1317D11 (DDBJ:  
AP006186) clone at 5' end and with OJ1567 (DDBJ: AP003759) at 3'  
end. Detailed information on overlap and assembly quality together  
with annotation of this entry is available at  
<http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

## FEATURES

## Source

1. .149371  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="7"  
/clone="P0585H11"

## gene

complement(298..792)

## mRNA

/genes="P0585H11.101"

complement(join(<298..591,655..>792))

/note="start and end point are not identified"

## CDS

complement(join(298..591,655..792))

/genes="P0585H11.101"

/notes="similar to Oryza sativa chromosome 1, P0409B08.8"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="BAC20734.1"

/db\_xref="GI:23617046"

/translation="WTAGRCFCAARRHGRQARAGAAAGDGGGQDAQVHHGTRARDATGG  
AAAAGQDRVETETARTAHGQPRGRANDARGAARERGELGSAPALAAHARGQS  
CAEGERAREKESERREGGDPRIPIRDPGGKIDFCGGI"  
complement(1507..1932)  
/gene="P0585H11.102"  
complement(<1507..>1932)  
/note="start and end point are not identified"  
complement(1507..1932)  
/gene="P0585H11.102"  
/note="predicted by GlimmerM etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC20735.1"  
/db\_xref="GI:23617047"  
/translation="MRRHAHQPHGRWVARTGCTDRPDPISAEVLPTWRLGAYVA  
ATRAGGRKTASAANGRAAASVGRLATMRGDGAYTMTKKRERANGLDSPGAR  
QRTISAATGGEQRKRRRHEGPIGGESIYAATGIRC"  
complement(5704..6661)  
/gene="P0585H11.104"  
complement(5704..6661)  
/gene="P0585H11.104"  
/note="supported by full-length cDNA(s): AK106062"  
5872..6872  
/gene="P0585H11.103"  
complement(5872..6872)  
/note="supported by full-length cDNA(s): AK109288"  
5872..6872  
/gene="P0585H11.103"  
/note="contains full-length cDNA(s): AK109288  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
complement(5942..6631)  
/gene="P0585H11.104"  
/note="contains EST(s): AU101213(E2609),C72987(E2609)  
contains full-length cDNA(s): AK106062"  
/codon\_start=1  
/product="putative disease resistance response  
protein-related/dirigent protein-related"  
/protein\_id="BAC20736.1"  
/db\_xref="GI:23617048"  
/translation="MAPTSPAPALFLALACAAAPLLRAADTHLHFFMHVWSGG  
QTAQVQIKGPTSAAGVSTGQDVTVDALTTETSSATSPVGRAGQYMMSSLSAPT  
LMCMVNLFTAGENNSTIAVGHDDTTATRELVSVGTKFRTATGYVVKTSMS  
ASTGVFELDVYVTTNATTIDASAPVPLDGGSGSGSTAKSAGARQVGVWSACV  
GLVVALVGRGW"  
complement(8221..8754)  
/gene="P0585H11.105"  
complement(join(<8221..8529,8623..>8754))  
/gene="P0585H11.105"  
/note="start and end point are not identified"  
complement(join(8221..8529,8623..8754))  
/gene="P0585H11.105"  
/note="predicted by GeneMark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC83532.1"  
/db\_xref="GI:34395174"  
/translation="MAVTGVGVLDGGWRWQACASRTWNRISGGDGLLLWSRWYVE  
ATAGCTPWSRPOCLSRVNSLENISLRPVAHRRFNSITAAAEERAPILGSL  
WNGVCTGHDLVVSSCMPLPQSSSSQDQLICSHNAPMLKVG"  
complement(14091..14753)  
/gene="P0585H11.106"  
complement(<14091..>14753)  
/gene="P0585H11.106"  
/note="start and end point are not identified"  
complement(14091..14753)  
/gene="P0585H11.106"  
/codon\_start=1  
/product="putative disease resistance response  
protein-related/dirigent protein-related"  
/protein\_id="BAC20739.1"



```

||||| 1 GATGAACCGCAATACGATTCTTGGGAGCAACATACTGTTCAGATCCGGGAGCAAAAT 60
||||| 542 GTTGTTCGGGAATGACAAACGGTCTGTATCTGGGAATAACAACCATGTGTCTGGGAGC 601
||||| 61 GTTCTTGTCTGGGAATGACAAACCGCTCATATCTGGGGACAAACAATAGTGTGTCTGGGAGC 120
||||| 602 AACACACTGTGTAACTGGAAGTCACAATCTGTAGTTGGTAGCAACCAATCTGTATCA 661
||||| 121 AACACACTGTCTGAAGTGGGAATGACAATACCGTAACCGGAGCAACATGTCTGTATCA 180
||||| 662 GGGCAAAAGCATATTTACTGATACCAATATTTGTATCCGGGAACGACAATAATGTG 721
||||| 181 GGGCAAAACCATATCGTTACAGACAAACAATACGTATCCGGGAACGATAAATGTA 240
||||| 722 TCTGGAAGTTCATACCTGTATCAGGGAGGACAAATACCGTATCCGGGAGCAACAATACT 781
||||| 241 TCCGGAGCTTTCATACCGTATCCGGGGGACAAATCTGTCTCGGGAGCAACAATACC 300
||||| 782 GTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTAAACAGATGGTTAA 838
||||| 301 GTATCTGGGAGCAACCATGTTGTATCTGGAAGCAACAAGTCGTGACAGACGCTTAA 357

RESULT 2
AX019971
LOCUS AX019971 357 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 2 from Patent WO9937782.
ACCESSION AX019971
VERSION AX019971.1 GI:10043803
KEYWORDS
SOURCE Loliium perenne
ORGANISM Loliium perenne
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Poaceae; Loliium.
REFERENCE
1 Twigg,S., Worrall,D., Jarman,C.D. and Sidebottom,C.M.
AUTHORS Twigg,S., Worrall,D., Jarman,C.D. and Sidebottom,C.M.
JOURNAL Frozen food product
TITLE Patent: WO 9937782-A 2 29-JUL-1999;
JOURNAL TWIGG SARAH (GB); UNILEVER PLC (GB); WORRALL DAWN (GB); JARMAN CARL
DUDLEY (GB); SIDEBOTTOM CHRISTOPHER MICHAEL (GB); UNILEVER NV (NL)
FEATURES
source
Location/Qualifiers
1..357
/organism="Loliium perenne"
/mol_type="unassigned DNA"
/db_xref="taxon:4522"

```

```

Query Match 27.0%; Score 264.2; DB 6; Length 357;
Best Local Similarity 83.8%; Pred. No. 2.4e-60;
Matches 299; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 482 GATGAAGAACCAATACAAATATCAGGACCAACAATAGTTGGATCAGGAGCAACAAT 541
DB 1 GATGAACCGCAATACGATTCTTGGGAGCAACAATCTGTTCAGATCCGGGAGCAAAAT 60
QY 542 GTTGTTCGGGAATGACAAACGGTCTGTATCTGGGAATAACAACCATGTGTCTGGGAGC 601
DB 61 GTTCTTGTCTGGGAATGACAAACCGTCTATCTCGGGACCAACAATAGTGTGTCTGGGAGC 120
QY 602 AACACACTGTGTAACTGGAAGTCACAATCTGTAGTTGGTAGCAACCAATGTCTGTATCA 661
DB 121 AACACACTGTCTGAAGTGGGAATGACAATACCGTAACCGGAGCAACATGTCTGTATCA 180
QY 662 GGGCAAAAGCATATTTACTGATACCAATATTTGTATCCGGGAACGACAATAATGTG 721
DB 181 GGGCAAAACCATATCGTTACAGACAAACAATACGTATCCGGGAACGATAAATGTA 240
QY 722 TCTGGAAGTTCATACCTGTATCAGGGAGGACAAATACCGTATCCGGGAGCAACAATACT 781
DB 241 TCCGGAGCTTTCATACCGTATCCGGGGGACAAATCTGTCTCGGGAGCAACAATACC 300

```

```

QY 782 GTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTAAACAGATGGTTAA 838
DB 301 GTATCTGGGAGCAACCATGTTGTATCTGGAAGCAACAAGTCGTGACAGACGCTTAA 357

RESULT 3
LPE277399
LOCUS LPE277399 357 bp mRNA linear PLN 29-APR-2000
DEFINITION Loliium perenne partial mRNA for ice recrystallisation inhibition
protein.
ACCESSION AJ277399
VERSION AJ277399.1 GI:7573547
KEYWORDS ice recrystallisation inhibition protein.
SOURCE Loliium perenne
ORGANISM Loliium perenne
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Poaceae; Loliium.
REFERENCE
1 Sidebottom,C.M.
AUTHORS Thesis (1999) University of York
JOURNAL 2 (bases 1 to 357)
REFERENCE Sidebottom,C.M.
AUTHORS Direct Submission
TITLE Submitted (11-APR-2000) Sidebottom C.M., Plant Science, Unilever
JOURNAL Research, Colworth House, Sharnbrook, Bedfordshire, MK44 1LQ,
UNITED KINGDOM
COMMENT cds represents presumed mature peptide generated by cleavage before
the first amino acid.
FEATURES
source
Location/Qualifiers
1..357
/organism="Loliium perenne"
/mol_type="mRNA"
/db_xref="taxon:4522"
<1..357
/function="freezing tolerance"
/note="ORF1"
/codon_start=1
/product="ice recrystallisation inhibition protein"
/protein_id="CAB87814.1"
/db_xref="GI:7573548"
/db_xref="UniProt/TrEMBL:Q9M3W4"
/translation="DEPNTISGNNVTVRSGSKNVLAGNDNTVISGDNNVSGSNNVT
VSGNDNTVSGNHVSGTNIHVTNNNNVSGNDNNVSGSFHTVSGGHNTVSGSNNTVS
GSNHVSGSKNVTD"
mat_peptide
1..354
/product="ice recrystallisation inhibition protein"
/note="ORF1"

```

```

ORIGIN
Query Match 27.0%; Score 264.2; DB 8; Length 357;
Best Local Similarity 83.8%; Pred. No. 2.4e-60;
Matches 299; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 482 GATGAAGAACCAATACAAATATCAGGACCAACAATAGTTGGATCAGGAGCAACAAT 541
DB 1 GATGAACCGCAATACGATTCTTGGGAGCAACAATCTGTTCAGATCCGGGAGCAAAAT 60
QY 542 GTTGTTCGGGAATGACAAACGGTCTGTATCTGGGAATAACAACCATGTGTCTGGGAGC 601
DB 61 GTTCTTGTCTGGGAATGACAAACCGTCTATCTCGGGACCAACAATAGTGTGTCTGGGAGC 120
QY 602 AACACACTGTGTAACTGGAAGTCACAATCTGTAGTTGGTAGCAACCAATGTCTGTATCA 661
DB 121 AACACACTGTCTGAAGTGGGAATGACAATACCGTAACCGGAGCAACATGTCTGTATCA 180
QY 662 GGGCAAAAGCATATTTACTGATAACAATATTTGTATCCGGGAACGACAATAATGTG 721
DB 181 GGGCAAAACCATATCGTTACAGACAAACAATACGTATCCGGGAACGATAAATGTA 240
QY 722 TCTGGAAGTTCATACCTGTATCAGGGAGGACAAATACCGTATCCGGGAGCAACAATACT 781
DB 241 TCCGGAGCTTTCATACCGTATCCGGGGGACAAATCTGTCTCGGGAGCAACAATACC 300

```

